ESSAYS ON THE ECONOMETRIC ANALYSIS OF DISTRIBUTIONAL EFFECTS WITH APPLICATIONS TO LABOR MARKET INSTITUTIONS IN GERMANY

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von

MARIAN VITUS RÜMMELE

Tübingen
2024

Tag der mündlichen Prüfung	02.12.2024
Dekanin:	Prof. Dr. Taiga Brahm
Dekan:	Prof. Dr. Dominik Papies
1. Gutachter:	Prof. Dr. Martin Biewen
2. Gutachter:	Prof. Dr. Joachim Grammig

Publication notes

Chapter 2 is a revised version of Biewen, M., B. Fitzenberger, and M. Rümmele (2022). Using Distribution Regression Difference-in-Differences to Evaluate the Effects of a Minimum Wage Introduction on the Distribution of Hourly Wages and Hours Worked. *IZA Discussion Paper No. 15534*.

Chapter 3 is based on Rümmele, M. (2024). Combining Difference-in-Differences and Recentered Influence Functions to Estimate Quantile Treatment Effects – Pitfalls and Remedies. *Unpublished manuscript, University of Tübingen*.

Chapter 4 is based on Rümmele, M. (2024). How Did De-unionization Impact the German Structure of Earnings? A Distributional Approach Using Grouped Quantile Regressions. *Unpublished manuscript, University of Tübingen*.

Funding

The research for this doctoral thesis was, in part, funded through the DFG Priority Program 1764 and DFG projects BI 767/3-1 and FI 692/16-1.

Danksagung

Die vorliegende Dissertation ist im Zeitraum meiner Anstellung als wissenschaftlicher Mitarbeiter am Lehrstuhl für Statistik, Ökonometrie und Quantitative Methoden an der Eberhard Karls Universität Tübingen entstanden. Die Fertigstellung dieser Arbeit wäre nicht möglich gewesen ohne die vielen Freunde, Familienmitglieder und Wegbegleiter, die Vertrauen in mich hatten und zu Zeiten an mich geglaubt haben, an denen es mir selbst schwerfiel, an mich zu glauben.

Zuallererst gilt mein Dank meinem Doktorvater Martin Biewen, der es mir nach meinem Masterstudium ermöglicht hat, die Promotionsreise antreten zu dürfen. Vielen Dank dir, Martin, für deine kontinuierliche Unterstützung, die hilfreichen Gespräche und die hervorragende Zusammenarbeit. Insbesondere während der intensiveren Phasen unseres gemeinsamen Projektes am Stala in Stuttgart und am IAB in Nürnberg war ich über die effiziente gemeinsame Arbeitsweise und deine Unterstützung sehr dankbar.

Auch meinem Zweitbetreuer, Joachim Grammig, möchte ich von Herzen danken für die vielen persönlichen und fachlichen Gespräche. Insbesondere während der "WG-Zeit" der beiden Ökonometrie-Lehrstühle in der Sigwartstraße war ich sehr froh darüber, dich als Zweitbetreuer gehabt zu haben. Vielen Dank für alles.

Eines der Dissertationskapitel ist in Zusammenarbeit mit Bernd Fitzenberger entstanden, dem ich an dieser Stelle sowohl für die Zusammenarbeit als auch für die Möglichkeit, die Analysen vor Ort am IAB durchführen zu können, herzlich danken möchte. In diesem Zusammenhang gilt mein Dank auch Miriam Schirmer und Karolin Hiesinger vom IAB.

Während meiner Zeit als wissenschaftlicher Mitarbeiter durfte ich zudem die vorliegenden Kapitel auf zahlreichen Konferenzen und Workshops vorstellen. Den Teilnehmern der "34th Annual Conference of the European Society for Population Economics" (2021), der Jahrestagung des Vereins für Socialpolitik (2021), der "Annual Conference of the European Economic Association" (2022), der Datennutzendenkonferenz der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder ("Forschen mit amtlichen Mikrodaten", 2022), der Annual Conference der "International Association for Applied Econometrics" (2023), des Tübingen Hohenheim Economics Christmas Workshop

(2023), der gemeinsamen Ökonometrie-Doktorandenseminare der Universitäten Hohenheim, Tübingen, Friedrichshafen, Konstanz und Freiburg (2021, 2023, 2024) sowie der "37th Annual Conference of the European Society for Population Economics" (2024) möchte ich herzlich für die vielen wertvollen Kommentare danken.

Meine Dissertationszeit war in großen Teilen gekennzeichnet von den Umständen, die mit der Covid-19-Pandemie einhergingen – umso wichtiger waren die vielen Freunde und Kollegen, ohne die diese Zeit um einiges belastender gewesen wäre, als sie es ohnehin schon war. Zu nennen sind hier insbesondere die Mitarbeiter der beiden Ökonometrie-Lehrstühle, die die Zeit meiner Promotion geprägt haben. Vielen Dank Lea Eiting, Alexandra Klotz und Sylvia Bürger für die unverzichtbare Hilfe bei administrativen Aufgaben. Vielen Dank an meine derzeitigen und ehemaligen Lehrstuhlkollegen Miriam Sturm, Pascal Erhardt und Jakob Schwerter. Besonders danken möchte ich Miriam und Jakob für die hervorragende Zusammenarbeit im Rahmen unseres gemeinsamen Projektes, das effizienter nicht hätte laufen können. Darüber hinaus möchte ich den vielen weiteren Wegbegleitern der Ökonometrie-WG für die zahlreichen gemeinsamen Mittagessen und Gespräche danken, die die Zeit für mich unvergesslich gemacht haben. Vielen Dank an Johannes Bleher, Dalia Elshiaty, Thomas Dimpfl, Stefan Glaisner, Alexander Reining, Julie Schnaitmann und Jantje Sönksen.

Darüber hinaus war meine Zeit in Tübingen geprägt von vielen Begegnungen und Freundschaften, ohne die die Fertigstellung dieser Dissertation so nicht möglich gewesen wäre. Zunächst bin ich sehr dankbar dafür, dass sich aus der anfänglich beengten Bürosituation mit Constantin Hanenberg eine tiefe Freundschaft entwickelt hat. Vielen Dank für die unvergessliche gemeinsame Promotionszeit mit dir. Ebenso gilt mein Dank Gregor Stahl, der mir seit Beginn meines Bachelorstudiums in Tübingen ein steter und wichtiger Wegbegleiter war und ist – vielen Dank für alles. Zudem möchte ich Georg Thunecke danken, der mich durch viele Höhen und Tiefen begleitet hat und mir seit dem Masterstudium immer mit Rat und Tat zur Seite stand.

Zu guter Letzt möchte ich meiner Familie danken, die mich seit Beginn meines Studiums unterstützt und begleitet hat. Insbesondere danke ich meinen Eltern, Manuela und Michael, dafür, dass ihr mir so viel freie Hand bei der Studienwahl gelassen und mich zu jeder Zeit in all meinen Entscheidungen unterstützt habt, sowie für euren unbedingten Rückhalt, auf den ich immer bauen konnte. Dies gilt gleichermaßen für Annkathrin, die mich mit ihrem Rückhalt und ihrer Unterstützung durch die Promotionszeit begleitet hat. Ohne deinen ansteckenden Optimismus und dein unerschütterliches Vertrauen in mich wäre vieles nicht möglich gewesen – dafür kann ich dir nicht genug danken.

Marian Rümmele, Tübingen im Juli 2024

Contents

List of Figures

List of Tables

Chapter 1

Dissertation Introduction

Few areas of economic research are as passionately debated outside of academia as labor market economics, primarily because virtually everyone is a market participant in one way or another. In other words, unlike other more ‘abstract’ markets, the labor market is deeply connected not only to measurable outcomes, but also to normative matters of economic justice, personal identity, and individual fulfillment. Given its pervasive role, it is worthwhile to ask: What are the features of the labor market that give it this distinctive position in public discourse and individuals’ world of experience? To ultimately answer this question, it is instructive to begin with a broad conceptualization of the driving forces behind observable labor market outcomes, distinguishing between a “competitive” and a “regulatory” institutional aspect of the market (Bound and Johnson, 1992; Autor and Katz, 1999). On the one hand, market-driven factors, including the supply and demand for various types of workers, labor-augmenting technologies, as well as globalization, play pivotal roles in determining wage rates for specific worker types, as well as in shaping the distribution of wages and earnings (cf., Acemoglu and Autor, 2011; Autor et al., 2013; Acemoglu and Restrepo, 2022). On the other hand, labor market institutions such as mandatory wage floors or collective bargaining play a decisive role in shaping outcome distributions as well (cf., Dustmann et al., 2009, 2014; Fortin et al., 2021).

The competitive side of the coin is arguably rather abstract and not immediately palpable – it is rarely the case that workers can directly assess whether the reason for their shifts in earnings or employment prospects are, e.g., technology-augmented shifts in demands for their skills. What can be experienced directly, however, is the effect of labor market institutions such as collective agreements or minimum wages. Put differently, the decline of unions or changes in minimum wage policies are directly linked to the reality of many people’s lives. This is also related to the notion that labor market institutions serve as a means to counteract potentially undesirable trends. Wages and earnings may be too (un)equally distributed for some, leading to public debate about how to counteract these developments. Ultimately, policymakers can respond to the demands of public discourse by either altering the statutory framework that governs labor market institutions or by establishing or abolishing institutions, such as minimum wage policies, themselves. In other words, the regulatory framework that emerges from these debates could create an environment that either strengthens or weakens labor market institutions (see, e.g., Salverda and Checchi, 2015, and the review of Saint-Paul, 2000, therein).

All of this underlines the fact that it is normative considerations, especially surrounding discussions on labor market institutions, that give the subject matter so much weight, extending beyond the academic realm. At the same time, these normative considerations must be linked to observable phenomena to connect intellectual conclusions with what can actually be experienced. That is, questions that concern 'high-level' theories of justice are inevitably intertwined with the empirical work that is done 'on the ground'. This notion is succinctly summarized by Green (2014, p. 293) in his attempt to highlight the role empirical labor market research should play within more fundamental normative considerations. There, he argues that

> "*(...) most standard theories of justice place a large weight on self- and social respect and (...) such respect has a lot to do with the position a person holds in the productive process – their wage and employment outcomes. That, in turn, means that assessments of justice in the real world hinge critically on how labour markets actually function in assigning wages and employment. The answers to these questions are ultimately empirical. This implies that political philosophers – to the extent that their theories of distributive justice are connected to the real world – should be interacting with empirical workers who study labour markets.*"

A rigorous quantitative assessment is, therefore, equally important for any normative assessment – without the former, the latter would be more or less poking around in the dark, or, as Green (2014, p. 303) puts it, "*(...) one cannot separate how labour markets function from decisions related to the justness of institutions and policies*".

Following this notion, this dissertation aims to contribute to the empirical foundation surrounding normative discussions on distributional effects in the labor market. That being said, the overall goal of this dissertation is to contribute to a better understanding of the impact of two of the most prominent labor market institutions in Germany, i.e., the recently introduced federal minimum wage as well as the impact of unions on the earnings structure, by examining how changes in these institutions affect the distribution of wages and earnings as well as wage and earnings inequality (chapters 2 and 4). In addition, chapter 3 of this dissertation discusses more generally the limitations and prospects of certain econometric tools for specific distributional analyses in the context of policy evaluation studies.

To set the stage for the subsequent chapters of this dissertation and to highlight the literature strands to which the respective chapters are intended to contribute, the remainder of this introductory chapter briefly outlines how previous contributions examined the effect of labor market institutions on labor market outcomes, focusing on their impact on wage and earnings inequality. Moreover, since German data are used in chapters 2 and 4, I also provide a brief overview of how wage and earnings inequality has evolved in post-reunification Germany, with an emphasis being placed on the role of labor market institutions. Finally, the chapters of the dissertation are briefly summarized.

Labor market institutions and earnings inequality

As described above, this dissertation revolves around questions that deal with the distributional effect of labor market institutions, namely minimum wage policies and the impact of unions in the labor market. To emphasize that these questions are ultimately empirical, it is important to note that, ex ante, the effect of neither minimum wage policies nor union wage setting is immediately obvious. On the one hand, mere economic reasoning points to both an inequality-enhancing and an inequality-reducing effect implied by union wage setting. The former arises because a union premium implies an increased wage wedge between more or less unionized sectors or firms, whereas the latter is implied by the fact that union wage setting is characterized by a more compressed wage distribution within a unionized firm or sector (Freeman, 1980; Card et al., 2004, 2020). On the other hand, the overall ex ante effect of minimum wage policies on the outcome distribution is not immediately obvious either. This is because a wage floor does not only induce a piling up of distributional mass at the minimum wage level but also implies spillover effects to wage groups above its level (compare, for example, Dickens and Manning, 2004; Manning, 2021).

The economics profession was not always concerned with (empirical) questions that touch on the topic of wage and earnings inequality. It was not until the 1980s that a debate about US wage inequality and the role it played in the labor market seriously took off (Salverda and Checchi, 2015). From the beginning, labor market institutions were seen as closely related to issues of the evolution of wage and income inequality (e.g., Freeman, 1980; Plotnick, 1982), even though the focus of the literature back then was more on ‘actual’ market mechanisms related to the supply and demand of labor as well as technology (see, e.g., review in Salverda and Checchi, 2015). However, under the impression of the stark increase of wage inequality in the US that coincided with a sharp decline in real minimum wage levels and a rapid decline in unionization in the 1980s, more and more studies in the 1990s focused on the institutional side of the labor market (Freeman, 1991; DiNardo et al., 1996; DiNardo and Lemieux, 1997, to name a few). Moreover, especially in light of the 'credibility revolution' of the 1990s and the corresponding increased interest in estimating the causal effects of the US minimum wage – as exemplified by Card (1992) and Card and Krueger (1994) – the literature on the institutional side of the labor market has grown massively.

Regarding empirical analyses of the impact of minimum wages on labor market outcomes, much of the public debate has, from the outset, focused on the potentially adverse employment effects of minimum wages (see Neumark and Shirley, 2022, for a detailed review of the US case). Nevertheless, the distributional effects of minimum wages have never been irrelevant (e.g., Card and Lemieux, 1996; DiNardo et al., 1996; DiNardo and Lemieux, 1997). The two strands of these early studies of the minimum wage policy in the US – the literature on the employment effect on the one side and the literature on the distributional effect on the other – could be viewed as somewhat walking in parallel. That is, there was not really a joint econometric framework that combined these two strands of the minimum wage literature. The employment literature was very much focused on difference-in-differences (DiD) designs (e.g., Card and Krueger, 1994), whereas the literature on the distributional effect was related to progresses in the field of econometric methodology (as for example in DiNardo et al., 1996, and their proposal of semi-parametric reweighting techniques).

In the years following these early studies and especially in more recent contributions, however, difference-in-differences techniques have become the overarching framework for both strands of the literature. Concretely, many recent studies employ distributional difference-in-differences approaches that allow for the study of a minimum wage effect on earnings inequality (e.g., Lee, 1999; Autor et al., 2016; Oka and Yamada, 2023; Xu et al., 2023), or on wage bins of the unconditional distribution, and thus enable one to draw a more comprehensive distributional picture (e.g., Cengiz et al., 2019; Gopalan et al., 2021; Cribb et al., 2021; Giupponi et al., 2024). Some other studies used unconditional quantile regressions (Firpo et al., 2009) in conjunction with linear difference-in-differences specifications to estimate the effect of the minimum wage along the quantiles of the unconditional distribution of earnings (Dube, 2019b; Gregory and Zierahn, 2022; Bossler and Schank, 2023; Caliendo et al., 2023; Bossler et al., 2024). The latter, however, comes with certain difficulties, as unconditional quantile regressions that make use of recentered influence functions (RIF) are not tailored for such treatment analyses and thus require the researcher to carefully investigate the method's appropriateness in the given context (see, for example, the remarks in Dube, 2019b).

Questions surrounding the appropriate econometric toolkit to estimate causal distributional effects play a major role in chapters 2 and 3. Chapter 2, on the one hand, presents our approach to providing a full distributional account of the minimum wage introduction in Germany. In doing so, we refrain from using the above-mentioned recentered influence function regression approach due to its inherent limitations and opt for a flexible combination of a distribution regression technique (Chernozhukov et al., 2013) and a difference-in-differences specification. On the other hand, chapter 3 delves into the identified difficulties of a 'RIF-DiD' approach and provides an in-depth examination of its inherent issues, while also proposing ways to overcome them.

As to the econometric possibilities for quantifying the distributional impact of the second labor market institution of interest in this dissertation, namely unions and collective bargaining coverage, the following provides a brief overview. A defining feature for studies that examine union wage setting is the institutional framework of the country under consideration (for an overview, see Bhuller et al., 2022). Ultimately, this framework defines what conclusions can be drawn from empirical analyses, where these conclusions are more limited, and which econometric tools can be sensibly employed in the first place. Broadly speaking, major differences across countries revolve around the level at which union-mediated wage setting takes effect, i.e., the level of the implied centralization. The institutional setting in the US and Canada, on the one hand, is characterized by an establishment-centered impact. This clear separation of a unionized and a non-unionized sector allows, e.g., for the possibility to use the (suitably adjusted) non-unionized sector as a counterfactual for the unionized sector to examine the distributional effect of union wage setting (Card et al., 2004, 2020). Furthermore, to point to an alternative method of leveraging the institutional framework, the phenomenon of 'union elections' in the US that take place within firms allows for the application of a regression discontinuity design (DiNardo and Lee, 2004) to estimate the effect of unions on wage outcomes. To summarize, being characterized by a fairly clear-cut distinction between a union and a non-union sector, the Anglo-Saxon system of industrial relation and collective bargaining essentially govern the econometric methods that can be employed to estimate the effect of unions on labor market parameters.

Similarly, the German system of industrial relations dictates the possibilities for empirical work using German data. Concretely, the German system is characterized by a higher level of centralization than, for example, the US system, rendering the mentioned empirical strategies unfeasible in the German context. This is because unions most prominently take effect through collective bargaining agreements that apply on the sector-region level, rather than on the establishment level (Fitzenberger and Sommerfeld, 2016). These agreements are, in turn, applicable to all establishments within the specified sector-region, provided that the respective employer is a member of the relevant employer association. Many studies aiming to quantify the effect of de-unionization in Germany focus on the extent to which the compositional shift in the workforce towards lower collective bargaining coverage explains the overall increase in inequality (e.g., Dustmann et al., 2009, 2014; Biewen and Seckler, 2019; Baumgarten et al., 2020). These contributions make use of reweighting techniques à la DiNardo et al. (1996) or hybrid RIF decompositions as described in Firpo et al. (2018), allowing for a clear separation of composition and wage structure effects. These studies also point out that the estimated results are not to be interpreted as general equilibrium effects, having to do with the German institutional setting. Concretely, the German system implies substantial spillover effects from firms that actively apply a collective agreement to 'uncovered' firms that are not legally bound by these agreements ('orientation without a formal contract', see detailed accounts in Bossler, 2019; Oberfichtner and Schnabel, 2019; Jäger et al., 2022). Taken together, research designs that exploit the establishment-centered system in the US and Canada are not feasible in the German context since the phenomenon of orientation implies "market spillovers or norms and expectations about pay" that obscure the distinction to a considerable extent (Jäger et al., 2022, p.17). With these caveats in mind, the study in chapter 4 builds on the fact that the major part of union wage setting in Germany occurs at the group – i.e., sector-region level – to characterize how shifts in collective bargaining coverage have affected the German structure of earnings as well as earnings inequality.

Evolution of wage and earnings inequality in Germany

Against the backdrop of rising wage and earnings inequality that was accompanied by the literature's increased interest in the topic of inequality in the Anglo-Saxon context, numerous contributions focused on the development and drivers of wage inequality in

Germany as well.[1] In the decade prior to German reunification, the evolution of wage inequality in Germany could be described as relatively stable, with some observable increases in upper tail inequality (Dustmann et al., 2009; Drechsel-Grau et al., 2022). This stands in contrast to the US (see, e.g., Autor et al., 2008; Firpo et al., 2018), where wage inequality rose steadily since the 1980s. The evolution of German wage inequality from the mid-1990s to the mid-2010s is characterized by comparable influencing factors that have been described for the increase in wage inequality in the US context a decade earlier (Dustmann et al., 2009). Therefore, similar to the US case, a substantial part of the literature emphasized the crucial role of labor market institutions in shaping wage and earnings inequality in post-reunification Germany (Dustmann et al., 2014; Biewen and Seckler, 2019; Ellguth and Kohaut, 2019; Jäger et al., 2022). Even though the phenomenon of deunionization occurred in both the US and Germany, the underlying reasons for the decline of unions were specific to the German context, as they resulted to a large extent from the developments that were brought about by the process of German reunification (Dustmann et al., 2014; Jäger et al., 2022).

In their account of the underlying reasons for the transformation from the 'Sick Man of Europe' to the 'Economic Superstar' that the German economy underwent from 1990 to 2008, Dustmann et al. (2014) highlight the stark increase in male earnings inequality as one of the most striking developments that took place in Germany. In bringing together the transformation that characterized the post-reunification years and impacted the evolution of wage inequality, Dustmann et al. (2014) particularly focus on the system of industrial relations in Germany. Given the economic challenges that the reunification process posed for the German economy, they emphasize that the decentralization of wage setting mechanisms in the form of de-unionization and the use of so-called opening clauses played a central role in counteracting the poor economic prospects at the end of the last century and the beginning of the current one. Especially the high imposed wage rates in the Western part of the country that were dictated by collective bargaining agreements could not be paid by Eastern employers, which prompted them to opt out of the system of centralized

[1] Note that the following overview of the evolution of wage and earnings inequality in Germany focuses on the role of labor market institutions. This implies that other crucial drivers, such as globalization (e.g., Dauth et al., 2014; Biewen and Seckler, 2019; Dauth et al., 2021), the role played by changes in employment patterns and part-time work (e.g., Biewen et al., 2018; Fitzenberger and Seidlitz, 2020), the detailed consideration of technological change and its impact on demands for worker types or tasks (e.g., Dustmann et al., 2009; Antonczyk et al., 2018), or approaches that take into account establishment-specific effects (e.g., Card et al., 2013), are deliberately neglected in this brief review.

bargaining. This development then also spilled over to West Germany, resulting in a decline in collective bargaining agreement coverage altogether (Card et al., 2013). This, in turn, came at the 'cost' of a sharp increase in wage inequality, with potential implications for poverty risk and social justice (Biewen and Juhasz, 2012). This point has been made by Jäger et al. (2022) as well, who note that the considerable decentralization of wage setting mechanisms since the 1990s has coincided with a more flexible labor market environment which fostered economic growth, but has, on the other hand, also been associated with a growing low-wage sector as well as rising wage and income inequality.

Taking into account several potential explanatory factors for the rise in male wage inequality between the mid-1990s and 2010, Biewen and Seckler (2019) conduct a detailed distributional decomposition analysis. They find that by far the most important compositional change that occurred during this period was the unprecedented de-unionization, explaining up to 60 percent of the distributional shifts that occurred during the period they consider. A comparable study by Baumgarten et al. (2020) also came to the conclusion that it was de-unionization that was one of the most important developments for the increase in wage inequality between 1996 and 2014. Building on the insight that collective bargaining is an important pillar of German industrial relations, chapter 4 provides a deeper look at the detailed and heterogeneous effect of union-mediated wage setting on earnings differentials across worker types and inequality in Germany. Moreover, the study in chapter 4 provides a nuanced characterization of how de-unionization between 1996 and 2014 has affected the German structure of earnings and inequality.

While wage inequality rose steadily until roughly 2010, this year also marks a turning point. The development up to this year was largely characterized by negative real wage growth rates at the lower percentiles of the wage distribution, moderate wage growth at the middle, and strong growth at upper percentiles (e.g., Dustmann et al., 2009, 2014; Fitzenberger and Seidlitz, 2020). It was also the development at the lower end of the distribution that was responsible for the fact that wage inequality came to a standstill in 2009/2010 and even declined somewhat in the following years (Felbermayr et al., 2016; Brüll and Gathmann, 2020; Drechsel-Grau et al., 2022; Bossler and Schank, 2023). This evolution was particularly strong in the Eastern part of the country as documented, for example, in Brüll and Gathmann (2020) who also examined how the developments in East Germany differ from the Western part of the country.

In particular, Brüll and Gathmann (2020) point to the importance of another labor market institution that played a major role especially in East Germany: minimum wage policies. While the federal minimum wage was only introduced in 2015, the first sectoral minimum wages had already been in place in some sectors since 1997 (see Fitzenberger and Doerr, 2016, for a more detailed review). Brüll and Gathmann (2020) highlight the importance of sectoral minimum wages in the Eastern part of the country, especially since these policies affected a much higher fraction of workers in East than in West Germany due to the comparatively lower wage levels. Consequently, while sectoral minimum wage policies were found to have had little impact in West Germany, the impact was found to have been substantial in East Germany. Concretely, regarding the distributional effect of sectoral minimum wages in East Germany, Gregory and Zierahn (2022) find that sectoral minimum wages played an important role in shaping wage distributions within affected industries.

Furthermore, against the backdrop of rising inequality that was largely characterized by persistently negative growth rates at the bottom of the wage distribution until 2010 that have been fueled by the unprecedented decline of collective bargaining coverage in Germany, the year 2015 marks a decisive turning point in German labor and social policy: The first-ever federally binding minimum wage was introduced. The introduction of a federal minimum wage at a rate of 8.50 euros per hour was not only a novelty in Germany but also represented a comparatively deep intervention, as the approximately 4 million employees affected by the policy could theoretically expect an average hourly wage increase of about 40% (Bossler and Schank, 2023). As stated by the then Minister of Labor and Social Affairs, Andrea Nahles, one of the main reasons for the introduction of the first federally binding minimum wage was the ongoing discussion about increasing inequality and the lack of wage floors for the majority of employees due to the dramatic fall in collective bargaining coverage in the years prior to 2015 (Nahles, 2014).

Although the introduction of the minimum wage was seen as a possible means of tackling the increased inequality, it was also accompanied by many voices fearing that this labor market intervention would lead to an increase in unemployment and thus ultimately fail to achieve its goal of reducing (earnings) inequality (e.g., Müller and Steiner, 2011; Knabe et al., 2014). After the introduction of the minimum wage, however, various evaluation studies found that fears of rampant unemployment as a result of the introduction of the minimum wage were exaggerated and that the introduction of the minimum wage was not accompanied by significant employment effects (Ahlfeldt et al., 2018; Caliendo et al., 2019; Bossler and Gerner, 2020; Bossler and Schank, 2023). Rather, in their widely acclaimed contribution, Dustmann et al. (2022) found that the introduction of the minimum wage led to reallocation effects away from less productive and towards more productive companies that were able to pay higher wages. Crucially, Dustmann et al. (2022) find that this shift was not accompanied by rising unemployment.

Regarding the distributional effect of minimum wage policies, a number of studies have recently been published that examine Anglo-Saxon countries (see above). However, there are still comparatively few publications that explicitly consider the effect of the newly introduced German minimum wage. Overall, the existing studies that examine the distributional effects of the minimum wage find significant distributional effects across the board. While some studies focus on the average wage growth of workers in wage groups at the bottom of the distribution (e.g., Ahlfeldt et al., 2018; Burauel et al., 2019), the study by Bossler and Schank (2023) is, at the time of writing this dissertation, the only published study that conducts a full distributional analysis of the newly introduced federal minimum wage. Importantly, they find that the effect of the minimum wage is not confined to the quantiles directly affected, but find evidence of substantial spillover effects up to the median of the distribution of monthly earnings. Until now, however, there has been no full account of the distributional effect on the hourly wage distribution – the distribution that is most directly affected by the minimum wage which constitutes an *hourly* wage floor. The study in chapter 2 aims to fill this gap in the literature.

Dissertation chapters

The last part of this introduction briefly summarizes the three subsequent chapters of this dissertation.

Using distribution regression difference-in-differences to evaluate the effects of a minimum wage introduction on the distributions of wages, hours, and earnings

In chapter 2, we provide a comprehensive account of the causal distributional effect of the introduction of the German national minimum wage. While previous studies have made use of small-scale survey data such as the German Socio-Economic Panel (Burauel et al., 2019; Caliendo et al., 2019, 2023) or were constrained to restrict their distributional analysis to monthly earnings using administrative social security data (Bossler and Schank, 2023), we make use of a large-scale database, the *German Structure of Earnings Survey*, that allows for an evaluation of the effect on the dimensions of hourly wages, monthly earnings, and hours worked. To take into account already existing trends prior to the introduction of the minimum wage, we further adjust our results by a pre-trend correction for which we make use of another large-scale database in the form of tailored administrative social security data provided by the Institute for Employment Research (*IAB*). Crucially, unlike in other years, these data also comprise hours worked for the years 2011 to 2014, allowing for an examination of existing pre-trends in all three dimensions we consider. Methodologically, we combine a distribution regression approach (Chernozhukov et al., 2013) with a difference-in-differences approach that makes use of 'bite' measures (Card, 1992) to trace the causal effect along the support of the respective dimension under consideration. Notably, this allows us to estimate the causal effect along the support of the *unconditional* outcome distribution as well as to derive the implied causal effect of the policy introduction on inequality measures. By using a distribution regression approach, we circumvent certain methodological difficulties of combining unconditional quantile regressions (Firpo et al., 2009) and difference-in-differences approaches. We furthermore show that our approach leads to an identification result that is similar to the results in Roth and Sant'Anna (2023), who formally derived identification conditions for distributional treatment effects. Our results show that the minimum wage did not only shift distributional mass from below the minimum wage threshold to the threshold but also spilled over to wage bins above it. While we also find significant induced shifts in the monthly earnings distribution, we find only small effects regarding hours worked for low-wage workers. Overall, our findings suggest that the implied hourly wage increases translated into changes in monthly earnings as well. Moreover, we find that the minimum wage introduction explains the bulk of the recent fall in wage and earnings inequality.

Combining difference-in-differences and recentered influence functions to estimate quantile treatment effects – pitfalls and remedies

Chapter 3 revolves around the above-mentioned difficulty of combining an unconditional quantile regression in the form of recentered influence function regression (Firpo et al., 2009) with a difference-in-differences approach to estimate unconditional treatment effects. Against the backdrop that several recent contributions (Havnes and Mogstad, 2015; Gregory and Zierahn, 2022; Caliendo et al., 2023; Bossler and Schank, 2023; Bossler et al., 2024) make use of unconditional quantile regressions in conjunction with linear difference-in-differences regression specifications, chapter 3 provides further insights into the conceptual limitations that occur even in the simplest case that involves two points in time and two groups. The structural shortcoming of a 'pooled RIF-DiD' approach – the approach that is employed in most of the aforementioned contributions – is due to the fact that recentered influence function approaches can only provide a first-order approximation of the actual effect. This requires the researcher to specify exactly around which quantile of the outcome distribution the effect is to be approximated. When pooling together both pre- and post-treatment distributions, it is shown that the estimated effect is not a reasonable approximation of the parameter of interest, i.e., the quantile treatment effect on the treated. Moreover, to overcome this shortcoming, two alternative approaches are proposed that, on the one hand, approximate around the correct quantile of interest ('DDiD-RIF'), and on the other hand, result in a discrete comparison of involved sub-populations ('QDiD-RIF'). These two proposed approaches are furthermore clearly linked to the required identification assumptions that were formulated in previous contributions (Athey and Imbens, 2002, 2006; Roth and Sant'Anna, 2023). A Monte Carlo exercise shows that the two proposed approaches perform much better than the pooled RIF approach if the underlying identification assumptions are met. To examine the performance outside this controlled Monte Carlo environment, the distributional impact of a large-scale redistributive US policy, the Earned Income Tax Credit (EITC), on income-to-poverty thresholds as in Hoynes and Patel (2018) is reexamined using all three approaches. This application suggests that differences between the two proposed approaches and the pooled RIF-DiD estimators for the quantile treatment effects on the treated are non-negligible in practice as well.

How did de-unionization impact the German structure of earnings? A distributional approach using grouped quantile regressions

Lastly, chapter 4 provides a nuanced characterization of the impact of unions in terms of collective bargaining coverage on the structure of earnings in Germany. In particular, the effect of the dramatic decline of collective bargaining coverage in Germany between the mid-1990s and 2014 on both between- and within-inequality measures, as well as on the unconditional distribution of earnings is examined. Methodologically, a novel approach that accounts for the specific German institutional setting of collective bargaining at the sector-region level in the form of a grouped quantile regression analysis (Chetverikov et al., 2016) is employed, allowing to examine several sources of effect heterogeneity across worker characteristics and the distribution of worker productivity. Following the approach outlined in Oka and Yamada (2023) and Xu et al. (2023), the empirical analysis highlights that the prevalence of collective agreements at the sector-region level affects worker-type specific earnings differentials, with the strongest effects being found for differentials between male and female workers and between workers with medium and low educational attainment. Using these results, it is furthermore demonstrated that the de-unionization in Germany between 1996 and 2014 accounts for a substantial fraction of the increase in both earnings differentials between as well as for the increase in inequality within worker types. Using the notion to simulate unconditional distributions from conditional quantile regression models formulated in Machado and Mata (2005), it is shown that the effect of de-unionization on the earnings structure clearly affected the unconditional distribution of real daily earnings as well, with the strongest effect at the lower-middle part of the unconditional distribution.

Chapter 2

Using Distribution Regression Difference-in-Differences to Evaluate the Effects of a Minimum Wage Introduction on the Distributions of Wages, Hours, and Earnings*

*This chapter is a revised version of Biewen, M., B. Fitzenberger, and M. Rümmele (2022). Using Distribution Regression Difference-in-Differences to Evaluate the Effects of a Minimum Wage Introduction on the Distribution of Hourly Wages and Hours Worked. *IZA Discussion Paper No. 15534*.

2.1 Introduction

Against the backdrop of a stark increase in wage inequality from the mid-1990s onwards (Dustmann et al., 2009; Antonczyk et al., 2010; Card et al., 2013; Biewen and Seckler, 2019), a national statutory minimum wage was introduced in Germany in 2015. The introduction of the German minimum wage at the level of 8.50 euros per hour constituted a major policy experiment: over 4 million workers (roughly 11% of the workforce) were eligible for it (Mindestlohnkommission, 2020).[1] Although there have been a number of recent contributions on the effects of the German minimum wage (Caliendo et al., 2018, 2019; Burauel et al., 2019; Dustmann et al., 2022; Bossler and Schank, 2023, and literature review below), the causal effect on the actual distribution of hourly wages and hours worked is an open question. The main aim of a minimum wage is to shift distributional mass from below to its level, leading to a spike of the wage distribution at the minimum wage. However, because of potential spillovers, its effects on the wage distribution may go beyond also shifting the wage distribution above the minimum wage (Brochu et al., 2023). A key challenge is to separate the causal effect of the minimum wage on the wage distribution from changes that would have happened anyway, i.e., from trends in wage setting policies of employers, in labor supply, and in wage bargaining, which may have started before the minimum wage introduction.

This study makes the following contributions. First, while there exists a considerable literature on the effects of the German minimum wage on various outcomes (see literature review below), this is the first study to make use of the scarce information on hourly wages and working hours in Germany from large-scale administrative databases. This allows us to reliably separate the effects of the minimum wage introduction along the distribution on prices (= hourly wages) from those on quantities (= hours worked). Minimum wages may potentially not only change hourly wages but also working hours. For example, firms may reduce working hours for low-wage employees to keep overall wage bills constant (Stewart and Swaffield, 2008), or there may be shifts between part-time and full-time employment as a response because the minimum wage may change the relative price

[1] Even though there existed a number of sector-specific minimum wages before (Fitzenberger and Sommerfeld, 2016), Germany was one of the few countries without a national minimum wage in the years prior to 2015. See Caliendo et al. (2019) for a comprehensive overview of research on the German minimum wage and its institutional details.

between the two (Garloff, 2019).

We use an innovative two-sample strategy to combine data from the *German Structure of Earnings Survey (GSES)* – which is the only German large-scale database that includes information on hourly wages and working hours both before and after the minimum wage introduction – and from the administrative *Deutsche Gesetzliche Unfallversicherung (DGUV-IAB)* database, which includes information on wages and working hours, but only for a few years before the minimum wage introduction. Both databases are considered highly reliable because a firms' participation is compulsory and the information on wages and hours is typically based on the firm's internal accounting system. In contrast, the studies by Caliendo et al. (2018), Burauel et al. (2019), Burauel et al. (2020), and Caliendo et al. (2023), which find reductions in hours worked in response to the minimum wage introduction, use survey data possibly suffering from relatively small sample size and potentially large measurement error in self-reported wages and working hours, thus possibly leading to noisy estimates and spurious findings of spillovers and noncompliance (Autor et al., 2016). Our contribution, using GSES and DGUV-IAB data, complements evidence from Bossler and Schank (2023) and Dustmann et al. (2022) solely based on the DGUV-IAB data, which, after the minimum wage introduction, do not include individual information on hours worked and thus not on hourly wages. Only reliable data on hourly wages allow to assess whether hourly wages are increased to the level of the minimum wage and whether there are spillover effects above. Evidence on hours worked reveals responses of firms at the intensive margin which may rationalize possible differences between the effects on hourly wages and earnings as suggested by the aforementioned studies using survey data.

As a second contribution, we develop a distribution regression difference-in-differences approach (DR-DiD) that may be of independent interest for many other applications in which the evaluation of distributional effects is important. A small number of previous contributions have carried out calculations related to the ones we present below (Almond et al., 2011; Dube, 2019b; Cengiz et al., 2019), but the previous literature has not produced a full statement of the approach for the whole distribution along with its identifying assumptions. In particular, we show that tackling the problem with a distribution regression naturally leads to an identification condition for distributional treatment effects recently shown by Roth and Sant'Anna (2023) to be equivalent to a parallel-trends assumption being independent of the functional form of the outcome variable. The distribution regression approach (Chernozhukov et al., 2013) appears particularly suited to study effects of the minimum wage on the distribution of hourly wages and hours worked as it directly targets nominal values instead of quantiles. It can also deal with discrete mass points and discontinuous distributions which may pose a problem for methods based on continuous distributions such as the Recentered Influence Function (RIF) regression (Firpo et al., 2009, 2018).

As a final contribution, we explore the use of alternative bite measures (based on regions, occupations and industries, respectively) as treatment indicators for the minimum wage introduction. This allows us to assess the sensitivity of our findings with respect to alternative channels for the minimum wage effect, e.g., concerning spillover effects, and potential violations of the no-pre-trends assumption.

The remainder of this paper is structured as follows. Section 2.2 provides a brief literature review. In section 2.3 and 2.4, we describe our data and econometric method. Section 2.5 presents empirical results, while section 2.6 concludes.

2.2 Related literature

A large literature analyzes the effects of minimum wages (e.g., Neumark and Wascher, 2008). In the following, we provide a selective review of contributions dealing specifically with minimum wage effects on the wage distribution, wage inequality, and hours worked.

A seminal contribution aimed at distributional effects of minimum wages is DiNardo et al. (1996). They used a 'tail-pasting' approach to construct counterfactual wage distributions in the absence of the minimum wage for the US from 1973 to 1992. The 'tail-pasting' approach rules out spillover effects of the minimum wage, evidence for which was found in an important contribution by Lee (1999). Lee (1999) exploited between-states variability in the minimum wage 'bite' in order to describe its effects on wage levels far above its threshold. His findings were later challenged by Autor et al. (2016) who used an instrumental variables approach to suggest that the spillover effects found by Lee (1999) might be 'measurement artifacts' stemming from imprecise wage and hours data. More recently, Cengiz et al. (2019) studied the impact of minimum wage changes on the wage distribution in the US. They find that minimum wage increases, which were amplified by modest spillover effects, boosted average earnings in low-wage jobs. Using the same method as Cengiz et al. (2019), Cribb et al. (2021) find that the introduction and subsequent increases of the UK National Living Wage from 2016 to 2019 led to substantial wage effects for workers at the lower tail of the distribution. Beyond this, the policy led to substantial spillover effects up to the 20th percentile, while no significant effects on employment were found. Based on reliable administrative payroll data, Gopalan et al. (2021) also find spillover effects up to 2.50 dollars above the minimum wage level accruing to incumbent as well as to newly hired workers, but only in firms with a significant fraction of low-wage workers. Building on DiNardo et al. (1996), Fortin et al. (2021) explicitly allow for spillover effects. They find significant evidence for spillover effects and show that allowing for spillovers substantially increases the contribution of minimum wage effects on changes in the wage distribution.

A number of contributions have analyzed the effects of the minimum wage introduction in Germany. An important general finding is the absence of significant employment effects (Caliendo et al., 2019; Dustmann et al., 2022; Bossler and Schank, 2023). Burauel et al. (2019) present evidence based on the German-Socio Economic Panel (GSOEP) suggesting excess hourly wage growth for low-wage workers. Also based on survey data from the GSOEP and using a regional bite measure, Caliendo et al. (2023) observe positive wage effects for the bottom hourly wage quintiles, but no significant effects on monthly earnings which they attribute to working hours reductions caused by the minimum wage, possibly reflecting noncompliance with the minimum wage by reducing paid hours but not actual hours. Based on the adminstrative DGUV-IAB data for all workers (full-time, part-time, marginal jobs), Dustmann et al. (2022) and Bossler and Schank (2023) have access to individual information on hours worked up to 2014 which allows them - like our study does - to define the regional bite of the minimum wage as treatment intensity. Dustmann et al. (2022) find that the minimum wage raised earnings in low-wage jobs and that reallocation to better paying firms accounting for around 17% of the earnings increases. Not having access to hours worked after the minimum wage introduction, the study divides earnings by the average hours worked in a labor market cell to proxy hourly wages. This approach does not allow to analyze the effect on the distribution of hourly wages. Moreover, Burauel et al. (2019), Dustmann et al. (2022) and Caliendo et al. (2023) focus on particular points in the distribution but do not provide a full distributional analysis aimed at measuring the impact of the minimum wage on the overall wage structure and wage inequality.

Using a Recentered Influence Function (RIF) approach modelling the quantiles of monthly earnings, Bossler and Schank (2023) provide a full distributional analysis for all workers. In contrast to Bossler and Schank (2023), our study focusses on the direct effect of the minimum wage on the distribution of hourly wages and hours worked which cannot be inferred from the administrative DGUV-IAB data after the minimum wage introduction. Estimating the effects on the distribution of hourly wages and hours worked adds to the exisiting evidence on the effects on the distribution of earnings. It yields further insights into the mechanisms behind the estimated minimum wage effects in Germany reported in the literature. Its effects should be visible in the distribution of hourly wages, whereas earnings effects also depend upon whether and to what extent there is an adjustment of hours worked at the intensive margin.

A smaller number of studies has focussed on the potential effects of the minimum wage on working hours. For example, Neumark et al. (2004) found that the U.S. minimum wage reduces hours worked for those paid at the minimum wage level with an elasticity of -0.3, but has no effect for workers receiving wages above the minimum wage. Stewart and Swaffield (2008) examined the effect of the British minimum wage on working hours and found a small total effect (including immediate as well as lagged effects) on weekly hours amounting to one to two hours per week. Dube (2019a) also found a small negative effect on working hours due to the introduction of the 2016 national living wage in the UK. For Germany, Burauel et al. (2020) find a significant decline in contractual working hours relative to unaffected workers but smaller and statistically insignificant effects on actual hours. Bachmann et al. (2020) present a comprehensive study of wage and hours effects of the minimum wage up to the year 2017 based on survey data (apart from the GSOEP, they exploit the so-called *Verdiensterhebung (VE)* which is similar in structure to the GSES but smaller and without compulsory participation). They conclude that there was a decline in hours in the year after the minimum wage introduction but find evidence that it was reversed later. Similarly, Bossler and Gerner (2020) exploit firm panel data to study, among other things, firms' behavioral responses to the introduction of the minimum wage. They find that firms reduced average working hours at the establishment level by 0.15 hours one year after its introduction (representing a 0.4 percent decrease in contractual working hours), but there were no significant shifts two years after its introduction. Taken together, the existing evidence on the effects of the German minimum wage on working hours is quite mixed, based on relatively small survey data and concentrates on the short-term effects in the first years after the introduction.

2.3 Data

The main part of our analysis is based on the *German Structure of Earnings Survey* (GSES) for the years 2014 (before the minimum wage introduction) and 2018 (after the minimum wage introduction). As mentioned above, the GSES is the only large-scale database for Germany that includes information on hours worked and thus hourly wages after the introduction of the minimum wage. The fact that the GSES is only carried out every four years makes an analysis of pre-trends difficult, especially given that there were major changes in the GSES sample design between 2014 and the preceding wave 2010. However, a pre-trend analysis is crucial for a credible difference-in-differences (DiD) analysis. We resort for this purpose to a specific administrative database from the *German Social Accident Insurance (DGUV)* containing information on individual working hours that can be merged with IAB data on employment histories for a few years before the minimum wage introduction (2011 to 2014), a unique dataset also used by Dustmann et al. (2022). Unfortunately, no such information on individual working hours is available from 2015 onward.

2.3.1 The German Structure of Earnings Survey (GSES)

We exploit the two most recent minimally anonymized waves of the GSES (2014 and 2018), which are only available on-site at the German statistical offices (see Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder, 2019). The GSES is a linked employer-employee dataset in which firms are legally obliged to participate and whose results are used for official statistical purposes. This ensures extremely low non-response rates of 2.3% in 2014 (Statistisches Bundesamt, 2016) and 3.2% in 2018 (Statistisches Bundesamt, 2020), respectively. The data included in the GSES can be considered highly accurate as most of them stem from firms' internal accounting systems which are transmitted electronically to the statistical agency. The GSES follows a two-stage sampling design. In the first stage, the statistical agencies draw from the full population of German firms (as listed in the official business registers). The second stage comprises the employees reported by a given firm, where the number of employees a firm has to report depends on the number of workers they employ. Sample weights ensuring the representativeness of the survey for the German dependent worker population are used by us throughout the analysis.

We impose a number of sample selection restrictions in order to address eligibility rules for the minimum wage as well as data limitations such as the missing regional information for particular groups of individuals (see appendix for details). Enforcing these sample selection restrictions yields our working sample covering 708,081 worker observations from 55,579 firms in 2014 and 693,827 worker observations from 55,722 firms in 2018, respectively.

2.3.2 Variables

Our earnings information is based on monthly gross earnings including overtime remuneration for the GSES reporting month April. Using data from April 2014 rules out possible anticipation effects of the newly introduced minimum wage.[2] Our data on hours worked refer to individuals' regular weekly working hours in the reporting month, including overtime hours. We follow the convention of transforming weekly working hours into monthly working hours by multiplying the former by the factor 4.345. The hourly wage measure is computed by dividing monthly gross earnings including remuneration for overtime hours by monthly hours worked including overtime hours. We do not adjust hourly wages by inflation as the minimum wage is likely to have an effect around its nominal level.

As individual characteristics, we consider sex, age, education, tenure, occupational position and occupation (KldB10, 2 digits). At the firm level, we include information on the federal state, individual information on remuneration according to collective agreements, firm size, whether the firm was part of the public sector, industry (WZ08, Statistisches

[2] Bossler and Schank (2023) use earnings information from employment spells that include June 30 of the calendar year considered, which includes additional earnings components typically paid to the employee during the second half of the calendar year. This explains why the earnings for April reported by the GSES are lower than the earnings averaged over a longer employment spell in the IAB, except for the 5%- and the 10%-quantile (see table A.3 in the appendix).

Bundesamt, 2008), as well as an indicator whether the firm was covered by a sectoral minimum wage (such sectors existed before the general minimum wage was introduced and continued to exist afterwards). The large size of our data set allows us to include all of this information in a very detailed way in our main analysis (see table A.2 in the appendix).

2.3.3 Bite measures

Our difference-in-differences approach relies on 'bite' measures reflecting the extent to which the minimum wage was going to affect certain subgroups of workers from the perspective of the pre-policy period. The seminal work by Card (1992) paved the way for a large body of contributions exploiting the bite measure derived from regional or other characteristics. The minimum wage bite in a particular population subgroup is defined as the fraction of individuals in this group with hourly wages below the minimum wage level before its introduction. This continuous group-level variation can be used to identify the effect of the minimum wage as wage adjustments are expected to be the stronger, the more workers in the respective group were below the minimum level before it was enacted. As the post-policy observation period is 2018, we compute the bite based on the minimum wage level of 8.84 euros/hour in that year (in 2017, the minimum wage was increased from the original level of 8.50 to 8.84 euros/hour). As a particular contribution, we use three different bite measures based on regions, occupations, and industries, respectively, based on pre-reform 2014 data. This allows us to investigate the sensitivity of the estimated minimum wage effects based on the treatment intensity as measured in different segments of the labor market, as there is no unambiguous measure of the strength of 'treatment' implied by the minimum wage introduction, and to explore different channels for spillover effects.

Bite 1: Local labor markets

A bite definition which has been used extensively in the literature is based on the relative impact of the minimum wage in different local labor markets. We use a definition of 96 German regions ('*Raumordnungsregionen*') as described in Bundesinstitut für Bau-, Stadt- und Raumforschung (2019).[3]

[3]Figure A.1 in the appendix provides an overview of minimum wage bites across regions.

Bite 2: Augmented occupations

An alternative bite measure is defined at the level of the occupations (e.g., Friedrich, 2020). Given the obvious importance of East-West differences, we augment the categorization according to 2-digit occupation codes (KldB10) by the information of whether the person worked in East or in West Germany. This yields a total number of 72 different groups.

Bite 3: Augmented industries

Finally, we define a bite measure for differences in the exposure to the new minimum wage across finely defined industries (WZ08). As in the case of occupations, we augment this categorization by information on whether the given person worked in East or in West Germany. Our industry bite measure augmented with the East/West information comprises 146 different groups.

Our motivation for using alternative bite measures is as follows. Since Card (1992), the most widely used bite measure has been the regional one, the idea being that in a region with a high fraction of individuals below the minimum wage, all workers can be considered as being potentially 'treated'. This allows for spillover effects to wages above the minimum wage in the local labor market as the labor market segment is considered. Such spillovers can be rationalized both in a competitive labor market model (demand for higher skilled workers increases as lower skilled workers become more expensive, see Neumark and Wascher, 2008) and in a monopsonistic model (higher wage workers are more likely to quit or do not start if wages are not raised by employers, altering the labor supply curve the monopsonistic employer faces, see Manning, 2003). Spillovers are particularly plausible under local monopsonistic competition, because the behavior of workers and employers depends strongest on the outside options offered by rival firms in the same region, whereas outside options in other regions would have to be sufficiently attractive to justify incurring the costs of regional mobility (Bhaskar et al., 2002; Bassier, 2021; Datta, 2021; Ransom, 2022). An additional explanation for spillovers to higher wage rates are fairness concerns, i.e., firms maintain wage differentials to prevent quitting (Dube et al., 2019). This particularly applies to the regional level, where workers can observe each other and also compete in other markets (e.g., housing). Since the minimum wage introduction reduces the monopsony power, one can expect potentially large spillover effects when using a regional bite measure.

In addition to neglecting spillover effects to less treated regions (e.g., by reallocation effects as found in Dustmann et al., 2022), a disadvantage of a very broad regional bite measure is that it may miss minimum wage effects for strongly exposed subgroups. Suppose, for example, that regional differences in exposure to the minimum wage are small. Nevertheless, it may be the case that individual subgroups such as certain occupations are strongly affected by the minimum wage. Defining bite measures at the level of, say, occupations, is appealing because of anecdotal evidence of pay shifts in certain occupations following the introduction of the minimum wage (hairdressers, cleaners, waiters etc.). This strategy follows the intuitive approach of studying to what extent wages changed differentially in occupations that were affected to a higher or lesser extent by the minimum wage. Given a potential lack of explanatory power in the regional bite, these effects may not be picked up when using this bite. Indeed, our overview of bite measures given in table 2.1 reveals stark differences in explanatory variation provided by the alternative bite measures, with the regional bite providing the least amount.

In addition to the occupational bite, we consider a bite measure defined across industries. This also turns out to contain more explanatory variation than the bite measure defined across regions. An additional advantage of a bite measure defined at the industry level is that a large part of wage bargaining in Germany takes place at this level (Jäger et al., 2022). This means that, in addition to target industries that are particularly affected by the minimum wage, an industry bite will be able to pick up minimum wage spillovers within industries. Given the persistent labor market differences between East and West Germany, we augment both the occupation bite and the industry bite by accounting for the differences between these two parts of the country.

A final motivation for considering alternative bite measures is that different bite measures may be differentially susceptible to violating the no pre-trend assumption in our difference-in-difference analyses. In particular, our results suggest that idiosyncratic developments at the regional level may produce irregular pre-trend patterns in some cases, while pre-trends at the occupation or industry level appear more stable, adding credibility to our estimation approach. Taken together, we view the use of alternative bite measures as complementary, allowing for a more complete picture of the available evidence and to assess the sensitivity of our results with respect to different aspects of exposure to the newly introduced minimum wage.

Table 2.1: Bite descriptive statistics

	Bite 1 German regions	Bite 2 Occupations + East/West	Bite 3 Industry + East/West
# Groups	96	72	146
Minimum bite	0.056	0.010	0.004
Maximum bite	0.320	0.634	0.759
Average bite	0.128	0.128	0.128
Standard deviation	0.062	0.129	0.138

Source: GSES 2014, own calculations.

2.3.4 Supplementary database for pre-trend analysis

By coincidence, the working hours information typically recorded by the *German Social Accident Insurance (DGUV)* can be merged with administrative employment data (Beschäftigenhistorik, BeH) provided by the Institute for Employment Research (IAB) just for the years 2011 to 2014 and for no other years. These data were also used by Dustmann et al. (2022). We use a 3.75 % sample of the BeH that was augmented with this working hours information for our pre-trend analysis. With some exceptions (see appendix), the DGUV-IAB data include the same covariate information as we use in the GSES. After applying the same sample selection criteria as in the GSES, our DGUV-IAB working sample covers 642,738 worker observations in 2011, 817,770 worker observations in 2012, 824,770 worker observations in 2013, and 831,304 worker observations in 2014, respectively. The use of the DGUV-IAB working hours information requires some pre-processing steps (see Dustmann et al., 2022; vom Berge et al., 2023, and appendix). As the wage data in the administrative employment data are top-coded, we only consider monthly earnings up to 4,050 euros per month and hourly wages up to 30 euros/hour.

2.4 Econometric method

Our aim is to determine the causal effects of the minimum wage introduction on the distributions of hourly wages, hours worked, and monthly earnings. One possibility would be to estimate a difference-in-differences version of a recentered influence function (Firpo et al., 2009, 2018). Some existing contributions have used such an RIF-DiD approach, see Havnes and Mogstad (2015), Dube (2019b), and Bossler and Schank (2023). In contrast to the applications pursued in these contributions, a RIF-DiD approach would be ill-suited for an analysis of minimum wage effects on the distribution of hourly wages as the introduction of a minimum wage is likely to introduce discrete mass points around its threshold which is in conflict with the assumption of continuous distributions underlying the RIF approach.[4] Moreover, the RIF approach is most easily applied to quantities such as quantiles and quantile ratios rather than to an analysis of changes in *nominal wage levels* at which the minimum wage is targeted. Both arguments equally apply to the distribution of weekly working hours which is known to be highly discrete and discontinuous.[5]

In order to address these aspects, we use a distribution regression difference-in-differences approach (DR-DiD). The distribution regression approach (DR) as developed by Chernozhukov et al. (2013) models effects on conditional and unconditional cumulative distribution functions by applying binary regressions to a range of thresholds of an outcome. A small number of previous contributions have carried out calculations related to the ones we present below (Almond et al., 2011; Dube, 2019b; Cengiz et al., 2019), but the

[4]By contrast, Dube (2019b) considers minimum wage effects on the distribution of *family incomes*, while Bossler and Schank (2023) focus on the distribution of *monthly earnings*. Both distributions are close to be continuous as minimum wage earners are spread over wide regions in these distributions. Appendix A.3 includes a discussion of further differences between the RIF-DiD and the DR-DiD approach, which may be of interest when deciding which method is best suited for an application.

[5]The method proposed by Brochu et al. (2023) is an alternative to our distributional analysis of the impact of a minimum wage on the distribution of hourly wages. The method uses standard flexible econometric models for the specification of the hazard rate of a distribution (usually applied to duration analysis) to estimate the effect on the wage distribution. By using a flexible specification of the baseline hazard in the wage dimension and applying the model for discretized wage bins, the method allows to estimate the shifts in the distribution at various levels wages (below the minimum wage, at the minimum wage and slightly above, and at higher wage levels) similar to our method. The method is implemented by Brochu et al. (2023) to estimate 'triple-differences' estimates of the minimum wage effects relative around the minimum wage. The method provides a simple comprehensive model of the conditional wage distribution in a very flexible way. Our modelling approach is more flexible in modelling the impact of the covariates on the distribution function in a distinct way at each wage level. However, the two modelling approaches are not nested and a careful specification analysis would be necessary to investigate which model fits the data in a better way.

previous literature has not produced a full statement of the approach along with its identifying assumptions. In particular, we show in the appendix that viewing the problem as a distribution regression and applying standard difference-in-differences assumptions to all thresholds naturally leads to an identification condition for distributional treatment effects as recently pointed out by Roth and Sant'Anna (2023) as a characterization of the assumption that the parallel trends assumption on the outcome is insensitive to functional form. The statement of the problem as a distribution regression naturally allows for the estimation of possible pre-trends, which we correct for as described below in section 2.4.2.

The identification assumption used by us has also been employed in a recent contribution by Kim and Wooldridge (2024) who develop a framework for evaluating quantile treatment effects. By contrast, our approach targets nominal points in distributions. We also point out a recent paper by Fernández-Val et al. (2024) (subsequent to our work), which provides a more detailed theoretical analysis of DR-DiD. In particular, Fernández-Val et al. (2024) advocate the use of non-linear distribution regressions models for DR-DiD. We fully agree with the arguments put forward by Fernández-Val et al. (2024), but, for reasons explained in section 2.4.3, use simpler linear probability models in our application.

2.4.1 Distribution regression difference-in-differences

We estimate the causal effect of the minimum wage using the continuous treatment measure $Bite_g$ (the minimum wage bite in group g) by estimating a large set of linear probability models for the cumulative distribution function (cdf) of the variable of interest based on the DiD model

$$\begin{aligned}\mathbb{P}(y_{igt} \leq z \,|\, Bite_g, D_g, D_t, X_{igt}) &\equiv F(z \,|\, Bite_g, D_g, D_t, X_{igt}) \\ &= \alpha_z + D_g \gamma_z + \lambda_z D_t + \beta_z (Bite_g \times D_t) + X_{igt} \delta_z, \qquad (2.1)\end{aligned}$$

where y_{igt} represents the observed outcome of interest of individiual i in bite group g at time t. The outcomes of interest in our case are either hourly wages, hours worked, or monthly earnings. The values z refer to a fine set of thresholds in the outcome distribution. For the case of hourly wages, we define the set $z \in \mathscr{W}$ such that we obtain wage bins $[0;3.49],[3.50;4.49],\ldots,[48.50,49.49]$ (after rounding hourly wages to the next integer cent value). Equation (2.1) describes the fraction of individuals with characteristics $(Bite_g, D_g, D_t, X_{igt})$ whose wage is less than or equal to threshold z. For the case of weekly hours worked, we first round hours to the largest integer below or equal and then define thresholds such that we obtain eight hours categories $[0;6],[7,11],\ldots,[42;50]$.[6] For monthly earnings we define an equally spaced set of thresholds ranging from 50 to 7,450 euros with a stepsize of 100 euros. For simplicity, the subsequent description of the econometric approach focuses on the hourly wage as the outcome of interest.

The variable D_g is a vector of dummies indicating to which bite group g individual i belongs. The term $D_g\gamma_z$ controls for time-constant differences in the fraction of individuals with hourly wages up to z between the different bite groups g. For example, if the bite is defined in terms of regions, D_g controls for the full set of regions. D_t indicates the pre-treatment ($t = 0$) and post-treatment period ($t = 1$), i.e., the term $\lambda_z D_t$ represents differences between periods 1 and 0 that are common to all individuals. Finally, we include a large set of observed characteristics X_{igt} which are also strong determinants of whether the observed wage does not exceed the threshold z. The characteristics considered are those shown in table A.2 in the appendix (naturally, for a given bite specification, the characteristic on which it is based, i.e., region, occupation or industry, is not included in X_{igt} as it is already included in D_g). In a sensitivity analysis, we will vary the set of characteristics X_{igt} used for conditioning, including the case in which we only specify the DiD terms $D_g\gamma_z, \lambda_z D_t$ and $\beta_z(Bite_g \times D_t)$, but no extra conditioning variables X_{igt}. The parameters in (2.1) are estimated by weighted least squares using the sample weights.

We model a linear impact of $Bite_g$ on the cdf of y_{igt}, i.e., β_z describes by how much the fraction of individuals below z was higher or lower in the treatment period $t = 1$ per unit of $Bite_g$ after controlling for all other observable characteristics. It is the part of changes that can solely be attributed to the degree of exposure to the newly introduced minimum wage but not to other determinants. The case $Bite_g = 0$ corresponds to the counterfactual situation with no minimum wage exposure. Consequently, the fraction of wages up to z

[6] In principle, it would be possible to define finer bins for working hours when using the GSES data only. However, it would be hard to connect such an analysis to potential pre-trends observed in the DGUV-IAB data whose working hours information is slightly different. In order to avoid potential misalignment, we define coarser hours categories.

in period 1 in the absence of the minimum wage is given by

$$F(z\,|\,Bite_g = 0, D_g, D_t = 1, X_{igt}) = F(z\,|\,Bite_g, D_g, D_t = 1, X_{igt}) - \beta_z Bite_g, \qquad (2.2)$$

i.e., the effects on the fraction of wages up to z, that are solely due to the differential exposure to the minimum wage, are subtracted.

Identification of this minimum wage effect is achieved under the assumption that $Bite_g$ is unrelated to factors influencing the wage distribution that are not captured by (D_g, D_t, X_{igt}) in our linear separable specification. In particular, there must not be differential time trends between groups g not captured by X_{igt}. This has to hold at each threshold z of the wage distribution. In section 2.4.2, we will investigate potential violations of this assumption in periods before the minimum wage introduction and use these observations to correct for pre-trends by augmenting (2.1) with a trend component estimated in the pre-period.

By the law of iterated expectations, the unconditional *factual* wage distribution in target period $t = 1$ is given by

$$F(z\,|\,D_t = 1) = \int F(z\,|\,Bite_g, D_g, D_t = 1, X_{igt})\; dF(Bite_g, D_g, X_{igt}\,|\,D_t = 1). \qquad (2.3)$$

By contrast, the unconditional *counterfactual* wage distribution in the absence of minimum wage effects is given by

$$F^{cf}(z\,|\,D_t = 1) = \int [F(z\,|\,Bite_g, D_g, D_t = 1, X_{igt}) - \beta_z Bite_g]\; dF(Bite_g, D_g, X_{igt}\,|\,D_t = 1). \qquad (2.4)$$

We show in the appendix how (2.4) is identified in repeated cross-sections under the assumption that standard parallel trends assumptions conditional on observables hold at each threshold z. This leads to the conditional analogue of an identification condition recently studied by Roth and Sant'Anna (2023), who show that this condition is equivalent to assuming that a parallel trends assumption is insensitive to functional form of the outcome and that the data generating process is a combination of random assignment and stationary potential outcomes. As we argue in more detail in the appendix, conditioning on a large number of observables and carefully addressing potential time effects (including those constructed from trends observed in pre-periods) make these conditions credible, thus securing the identification of the counterfactual distribution (2.4).

As cumulative distribution functions are more involved to interpret and in order to calculate inequality measures, we construct grouped probability functions based on the increments across the set of ordered thresholds $z \in \{z_0, z_1, \ldots, z_J\}$ defining J intervals $(z_{j-1}, z_j]$ $(j = 1, \ldots, J)$ by

$$f_{j,t} = F(z_j \,|\, D_t) - F(z_{j-1} \,|\, D_t), \tag{2.5}$$

$$f^{cf}_{j,1} = F^{cf}(z_j \,|\, D_t = 1) - F^{cf}(z_{j-1} \,|\, D_t = 1). \tag{2.6}$$

We use the following interpolation formulas for grouped data in order to calculate inequality measures and quantiles (Tillé and Langel, 2012). For the quantiles, this is

$$Q_t(\tau) = z_j + \frac{\tau - F(z_{j-1} \,|\, D_t)}{f_{j,t}} \, (z_j - z_{j-1}), \tag{2.7}$$

for τ such that $F(z_{j-1} \,|\, D_t) \leq \tau < F(z_j \,|\, D_t)$ and $t \in \{0, 1\}$. The one for the Gini coefficient is

$$Gini_t = \frac{1}{2\bar{z}} \frac{N_t}{N_t - 1} \sum_{j=1}^{J} \sum_{k=1}^{J} f_{j,t} f_{k,t} |z^c_j - z^c_k| + \frac{1}{\bar{z}} \sum_{j=1}^{J} \frac{\left(N_t f^2_{j,t} - f_{j,t}\right) L_{j,t}}{6(N_t - 1)}, \tag{2.8}$$

where N_t is the sample size, $z^c_j = (z_j + z_{j-1})/2$ the center of group j, $\bar{z} = \sum_{j=1}^{J} f_{j,t} z^c_j$ the group-implied estimator for the mean, and $L_j = z_j - z_{j-1}$ the length of the jth wage interval. For the right-open top group $j = J$, we make the following choices. Its length is chosen to be $L_{J,t} = z^{max} - z_{J-1}$, where z^{max} is the highest value observed in the sample. Its probability mass is given by $f_{J,t} = 1 - F(z_{J-1} \,|\, D_t)$ by the definition of the cdf. As the center of the last group, we always take the average value of y_{igt} in that group as observed in the factual distribution. Reassuringly, these formulae based on the group information lead to values that are very close to the ones coming from the usual nonparametric formulas.

The ceteris paribus effects of the minimum wage introduction on the distribution and on inequality measures are given by

$$\Delta_j^{cf} := f_{j,1} - f_{j,1}^{cf}, \quad j = 1, \ldots, J, \tag{2.9}$$

$$\Delta^{cf}(v(\cdot)) := v(F(z\,|\,D_t = 1)) - v(F^{cf}(z\,|\,D_t = 1)), \tag{2.10}$$

where $v(\cdot)$ denotes either quantiles or inequality measures (Gini and quantile ratios) computed from the full distribution.

2.4.2 Pre-trends: Estimation and Correction

The identification of the counterfactual wage distribution (2.4) is only valid if there are no other time trends in wages that differ across bite groups. For example, if the minimum wage bite is defined for regions, then it must not be the case that low-wage growth (conditional on covariates) was higher in high-bite than in low-bite regions as this would make the wage boosting effect of the minimum wage introduction appear higher than it was. To estimate potential differences in wage growth across different bite levels before the minimum wage introduction, we run regressions as in (2.1) for the pre-introduction period 2011 to 2014

$$\begin{aligned} F(z\,|\,Bite_g, D_g, year, X_{igt}) = \alpha_z + \sum_{t=2011}^{2014} \lambda_z^t \times \mathbb{1}[year = t] \\ + D_g \gamma_z + \sum_{t=2011}^{2014} \beta_z^t (Bite_g \times \mathbb{1}[year = t]) + X_{igt}\delta_z \end{aligned} \tag{2.11}$$

(compare Dobkin et al., 2018; Ahlfeldt et al., 2018; Freyaldenhoven et al., 2021, for the non-distributional case).

Here, we define the year $t = 2014$ as the reference period so that all coefficients concerning 2014 are normalized to zero (i.e., $\lambda^{2014} = 0, \beta_z^{2014} = 0$). The coefficients $\beta_z^{2011}, \beta_z^{2012}, \beta_z^{2013}$ represent systematic differences in wage growth for different levels of the minimum wage bite in pre-treatment years. The hypothesis of no pre-trends can be tested as $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$. If the coefficients $\beta_z^{2011}, \beta_z^{2012}, \beta_z^{2013}$ display systematic patterns (which they do in our application), we can extrapolate these patterns to the post-treatment period. For example, if the likelihood of falling under the hourly wage threshold of 8.5 euros/hour declined in high-bite regions in a systematic way faster than in low-bite regions before the minimum wage introduction, then one should subtract the extrapolation of this effect from the minimum wage effect in the post-period (because the fraction of wages below 8.5 euros would already have more strongly declined in these regions without the minimum wage). Section 2.5 discusses the estimated patterns of $\beta_z^{2011}, \beta_z^{2012}, \beta_z^{2013}$ in the pre-treatment period. For hourly wages, the pre-trends follow linear time trends almost exactly, which we then use for counterfactual trend extrapolation.[7] For hours of work and monthly earnings, the findings are more ambiguous with the pre-trends showing a less clear and often nonlinear pattern.

Formally, let $\bar{\Delta}_z$ denote the extrapolated effect of the pre-trend for wage threshold z. Then, the counterfactual wage distribution in the absence of the minimum wage *corrected for pre-trends* is given by

$$\begin{aligned} F^{cf,trend}(z\,|\,D_t=1) = \int & [F(z\,|\,Bite_g, D_g, D_t=1, X_{igt}) \\ & - \left(\beta_z - \bar{\Delta}_z\right) Bite_g\big]\; dF(Bite_g, D_g, X_{igt}\,|\,D_t=1), \end{aligned} \tag{2.12}$$

i.e., the extrapolation of the pre-trend has to be subtracted from the estimated effect of the bite. In section 2.5.1, we will consider different scenarios of extrapolating pre-trends, e.g., $\bar{\Delta}_z^1$ is the pre-trend effect under the assumption that the pre-trend lasts up to one year after the minimum wage introduction, and $\bar{\Delta}_z^2$ up to two years afterwards.

2.4.3 Estimation and specification

All factual and counterfactual distribution functions and their derivatives can be estimated by their sample counterparts (i.e., weighted sample averages using the sample weights). We compute bootstrap standard errors for all quantities based on clustering for the bite groups defining the treatment units (Bertrand et al., 2004).

Unfortunately, the data on which our analysis are based upon can only be accessed on-site in two separate research data centers. This has important consequences for our estimation strategy. First, we face substantial computational limitations on-site that make the use of

[7] The approach is analogous to, e.g., Dobkin et al. (2018), who only consider the part of the DiD-effect that deviates from a linearly extrapolated time trend.

the computationally more involved logit or probit models, as suggested by Fernández-Val et al. (2024), for (2.1) infeasible. Second, in a non-linear distribution regression model it would not easily be possible to combine our main analysis with a pre-trend analysis based on a *separate* data set, while this is straightforward in a linear probability model (see equation (2.12)). In general, we found in preliminary experiments with logit models that the factual and counterfactual *unconditional* distributions (2.3) and (2.4) were basically insensitive to the use of different models (linear proability vs. logit models) and/or covariate specification choices (inclusion/exclusion of covariates and interaction terms). This is unsurprising given the large amount of averaging involved. We point out the further practical advantages of linear probability models in the given context: (i) computational efficiency (the large number of parallel regressions for the different distributional thresholds can be efficiently parallelized as in Chernozhukov et al. (2020)), (ii) computational simplicity (avoidance of convergence problems for logit/probit models in case of near-perfect-prediction problems which may easily arise at extreme distributional thresholds), (iii) transparency, (iv) consistent aggregation due to the law of iterated projections, and (v) immediate interpretation of β_z in terms of percentage points probability mass gained/lost per unit of bite.

Still, one might be concerned that the 'rigid' nature of the linear probability model may lead to ill-defined treatment effects as described in Fernández-Val et al. (2024). Following Roth and Sant'Anna (2023) and Kim and Wooldridge (2024), we therefore subject our resulting counterfactual distributions (2.4) and (2.12) to the test whether they are proper distribution functions. Our tests do generally not yield evidence for specification problems with the exception of the lowest two thresholds in the specific case of hourly wages with two-year trend-correction (indicating trend over-correction in (2.12)). For details, see appendix A.2.

2.5 Empirical results

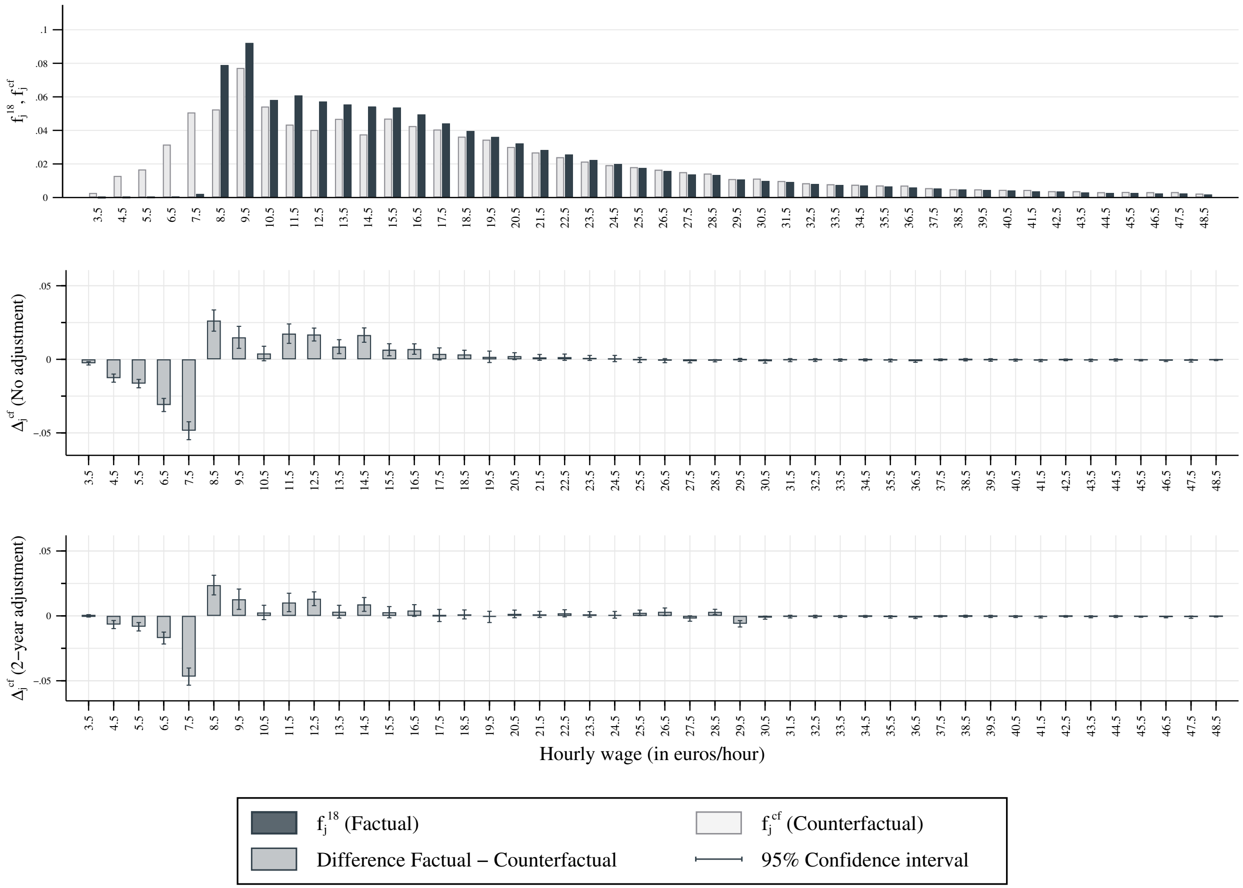
2.5.1 Effects on hourly wages

Figures 2.1 to 2.3 show the effects of the minimum wage on the distribution of hourly wages as measured by the three alternative bite definitions. The upper panels in each figure compare the factual distribution in 2018 with the counterfactual distribution that would have prevailed had the minimum wage not been introduced. The middle panels display the differences in the factual and the counterfactual frequencies of hourly wages in each bin of the upper panels as defined in eq. (2.9) without correcting for pre-trends. The lower panels display the differences when including the pre-trend correction.

The results based on the regional bite definition are presented in figure 2.1. The dark bars for the factual distribution in the upper panel suggest that the minimum wage was highly effective in eliminating hourly wages below its nominal level (8.84 euros/hour in 2018). The light bars in the upper panel of figure 2.1 depict the situation that would have prevailed under a hypothetical hourly wage structure without the minimum wage as inferred from the differential behavior of distributional change across regions. The differences between the factual and the counterfactual distribution shown in the middle panel visually demonstrates how the minimum wage shifted wages from below its level to wage bins above it. Apart from the fact that very low hourly wages were effectively eliminated, the results imply sizeable and significant spillover effects up to 16.5 euros/hour which is more than 80% above the nominal level of the minimum wage.[8] Also note the precisely measured zero effects for higher wage bins which can be interpreted as a validation check for our method because we would not expect causal effects of the minimum wage on very high wages.

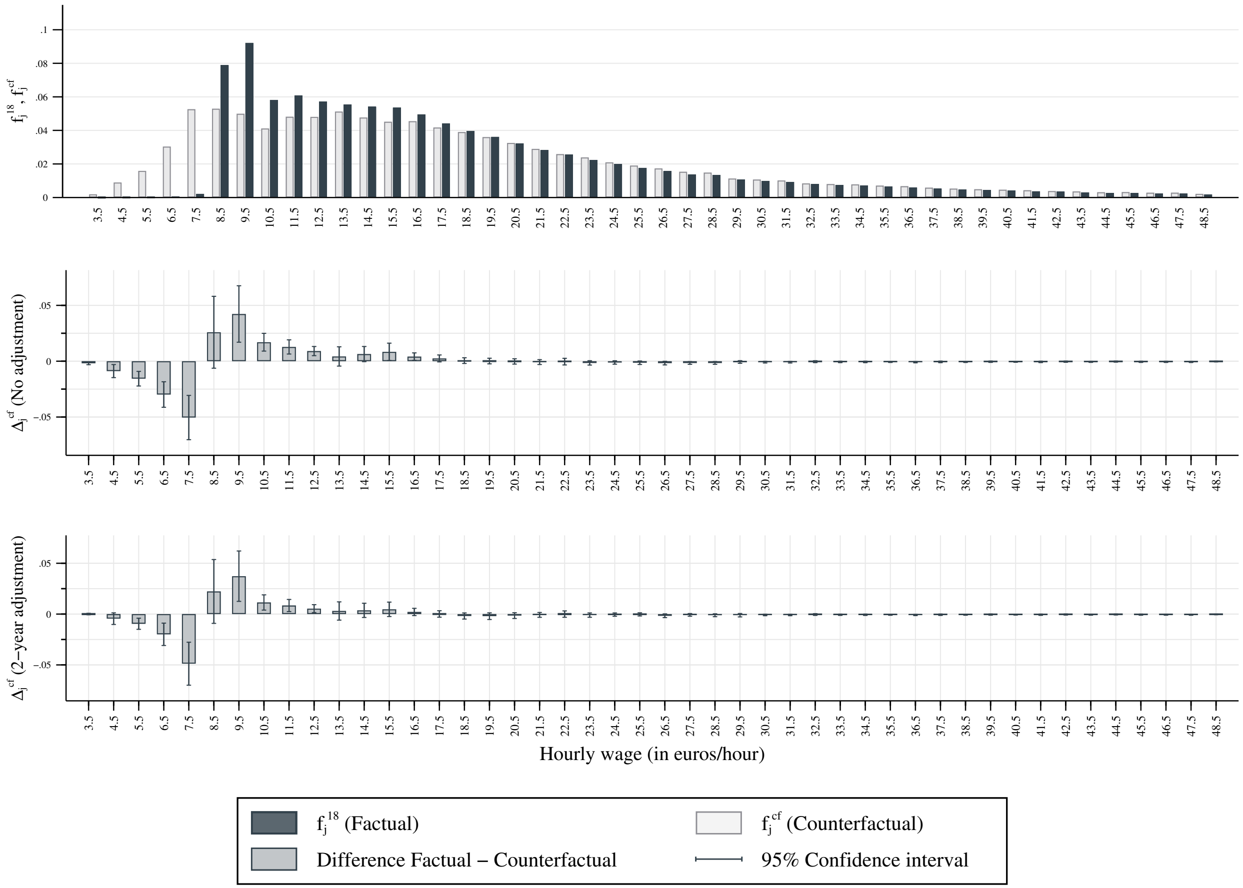
[8] We point out that some authors have noted that the GSES might overstate compliance with the minimum wage as employers may be hesitant to report non-compliant hourly wages, see Mindestlohnkommission (2020) and Bachmann et al. (2020).

Figure 2.1: 2018 Factual vs. counterfactual distribution of hourly wages in the absence of minimum wage.
Bite 1: Regions



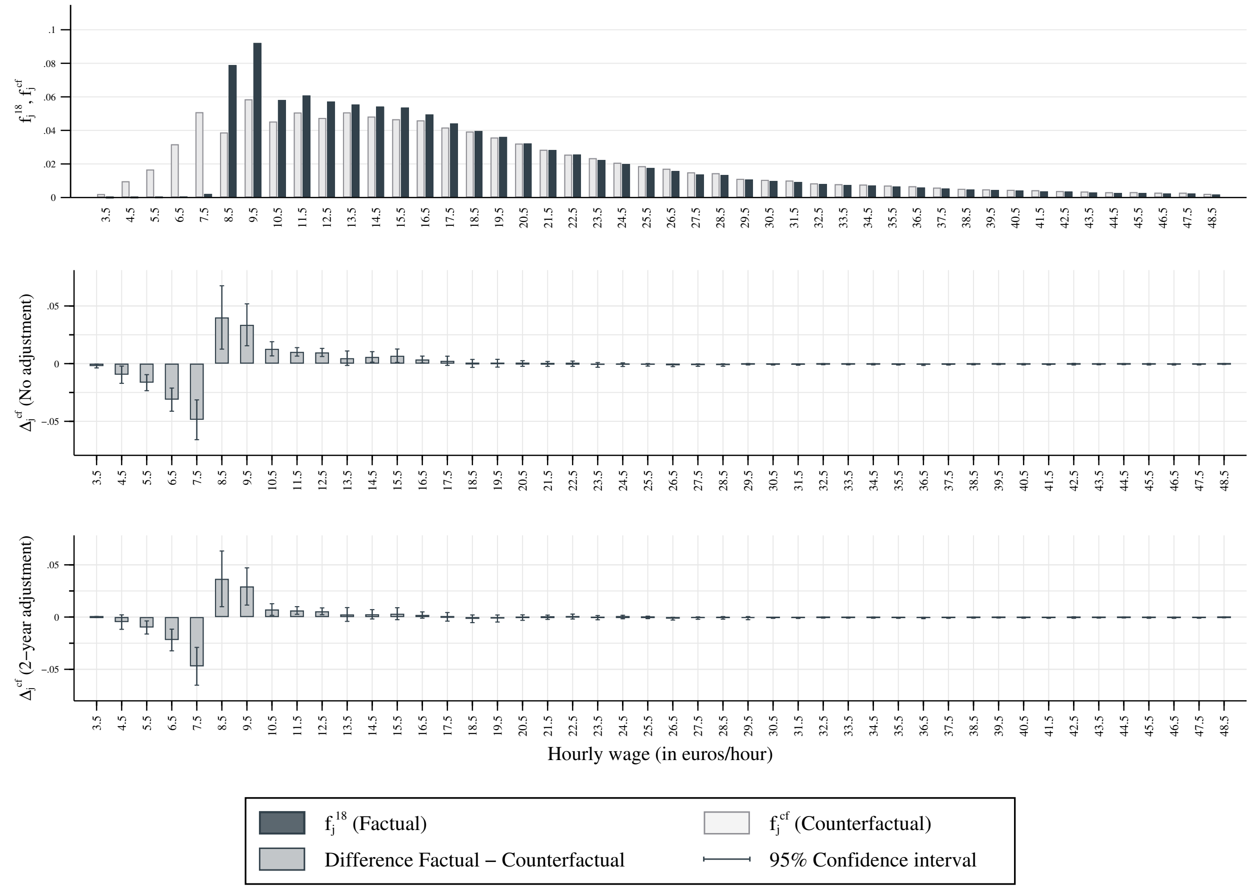
Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure 2.2: 2018 Factual vs. counterfactual distribution of hourly wages in the absence of minimum wage.
Bite 2: Augmented occupations



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure 2.3: 2018 Factual vs. counterfactual distribution of hourly wages in the absence of minimum wage. Bite 3: Augmented industries



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

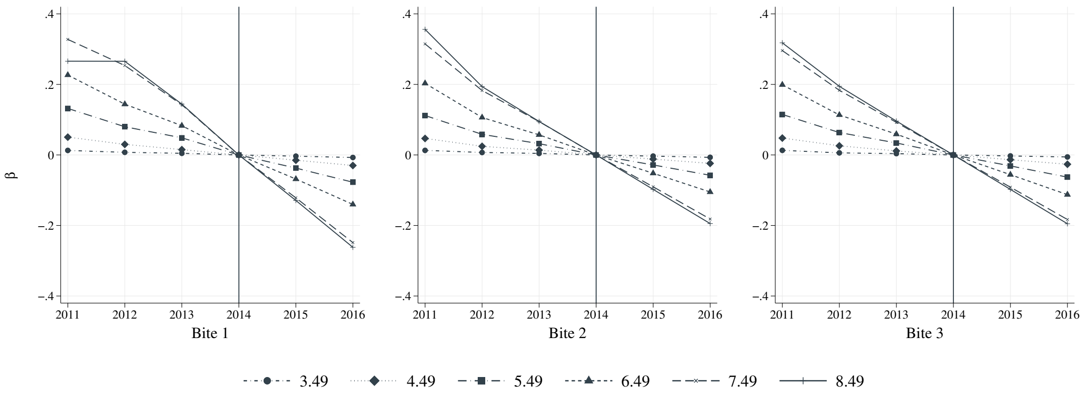
Figures 2.2 and 2.3 show the corresponding estimates based on bite differences across occupations (bite 2) or industries (bite 3), each augmented by the East/West distinction. The overall pattern looks quite similar to the one in figure 2.1, but significant spillover effects are less spread out, ranging only to 12.5 euros/hour (there are positive point estimates also for higher wage bins, but these are not or only marginally significant). The fact that measured spillover effects are larger for the regional bite suggests spillover effects within regions that are not picked up by the other two bite definitions.

The results presented above are not valid if there were differential time trends across the subgroups that define the bite variable in the years preceding the minimum wage introduction. For example, if the fraction of low wages had fallen more strongly in high-bite regions than in low-bite regions even before the minimum wage introduction, this trend would have been likely to continue after the minimum wage introduction. Then, one would incorrectly attribute part of the wage increases after 2015 in the lower tail of the distribution to the minimum wage. In this section, we demonstrate that such pre-trends indeed existed and we correct for them.

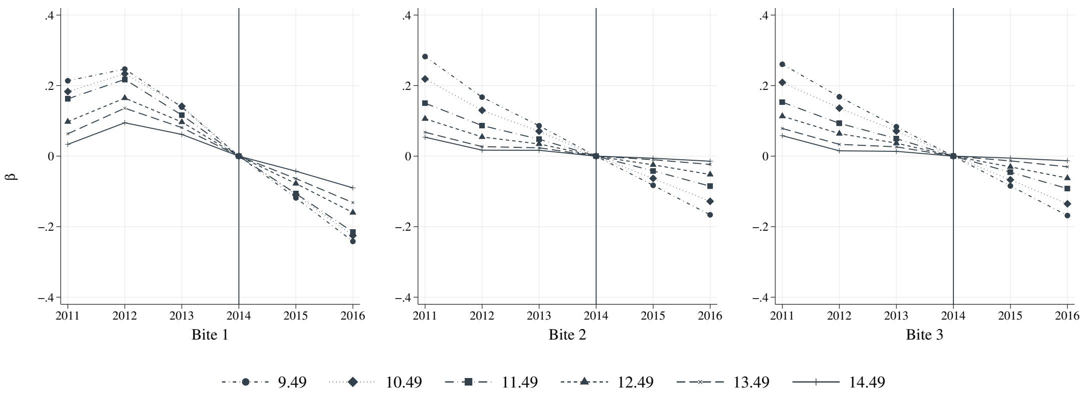
The estimates of the pre-introduction coefficients of the bite variable are shown in figure 2.4 (these are the $\hat{\beta}_z^t$ coefficients in equation (2.11)). In the absence of pre-trends, it should be the case that $\beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = \beta_z^{2014} = 0$, i.e., the trend in the likelihood for a wage up to z before the minimum wage introduction should not have been systematically different in high-bite compared to low-bite groups. Moreover, if the degree by which $\hat{\beta}_z^t$ differed from zero displayed a systematic trend in the years before the minimum wage introduction, this trend can be extrapolated to years after 2014.

For example, take the case of $z = 8.49$ euros/hour in the upper panel of figure 2.4 (solid line). In the years before the minimum wage introduction 2011 to 2013, individuals in high-bite regions were more likely to have wages below 8.49 euros/hour than in low-bite regions ($\hat{\beta}_z^t > 0$), but this was less and less the case, i.e., wages in high-bite regions already caught up to those in low-bite regions before the minimum wage introduction.

Figure 2.4: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Hourly wages, all bites.



(a) Lower thresholds (3.49 to 8.49)



(b) Upper thresholds (9.49 to 14.49)

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for bins below and above the minimum wage level. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

In the area right of the vertical bar, we extrapolate this trend linearly up to 2015 and 2016 (one year extrapolation and two year extrapolation). In a conservative approach focusing on the local behavior around the minimum wage introduction, we only use the years 2012 to 2014 to fit the pre-trend and extrapolate up to two years after 2014. The values of the extrapolated trend at 2015 and 2016 are therefore the values $\bar{\Delta}_z^1$ and $\bar{\Delta}_z^2$, which we have to subtract from the coefficient of the bite effect after 2014, because these represent by how much the fraction of wages up to z would have declined in high-bite compared to low-bite regions by the differential time trends alone (eq. (2.12)).

As figure 2.4 shows, there were systematic differential time trends years before the minimum wage introduction that are *uniform across all bite definitions*. Incidentally, the strength of the pre-trends is increasing constantly up to the 8.50 euros/hour threshold and then decreasing above. This means that the fraction of wages below 8.50 euros/hour was already declining more strongly in high-bite than in low-bite groups before the minimum wage introduction, indicating that the minimum wage effects may be overestimated without subtracting these effects. The fact that the observed patterns are uniform across the alternative bite definitions suggests that wage growth was already higher for low-wage workers in high-bite groups before the minimum wage introduction, independently of whether these bite-groups are defined by region, occupation, or industry. This points to exceptional wage growth for low-wage workers even in the years preceding the minimum wage introduction – and, incidentally, the effect was strongest around the minimum wage. Note that this pattern is not driven by East/West differences (detailed results are available upon request).

To what extent does the existence of these pre-trends change our estimated effects of the minimum wage? The lower panels of figures 2.1 to 2.3 show that the magnitude of the measured minimum wage effects is reduced by accounting for pre-trends, but only to a limited extent. A reason for the limited changes induced by the trend-correction is that the original distribution regression refers to the cumulative distribution function, while for the histogram bins the differences of the cumulative distribution function across adjacent thresholds matter (see equation (2.5)). As long as the trend-correction terms $\bar{\Delta}_z$ vary relatively smoothly across thresholds (as they do), their effect on histogram bins is limited. Still, we conclude that the impact of the minimum wage is somewhat overestimated if pre-trends are not accounted for.

The patterns in figures 2.4 can be interpreted as evidence against anticipation effects as there existed basically linear trends since at least the year 2012, which did not accelerate in the year 2014. For a discussion of potential anticipation effects of the German minimum wage, see Bossler (2017). As mentioned above, our GSES wage measure refers to April 2014. The parliament decided about the introduction of the minimum wage in July 2014 after intensive political debates earlier in the year. Recall, however, that the minimum wage did only come into force on January 1, 2015. Generally, it is unclear why employers should pay higher wages long before the introduction of a minimum wage if they are not obliged to do so (altruistic employers may always pay wages above the market level independently of a minimum wage). Based on IAB data for daily earnings averaged over employment spells until the end of 2014, Bossler and Schank (2023) also find little evidence for anticipation effects in 2014. Note that their data cover the whole year 2014, whereas we consider wages reported for April 2014 only.

How do these effects translate into changes of inequality measures? Asking this question is important as it is only in this way that one can assess the contribution of the minimum wage to general trends in wage inequality. Table 2.2 shows that wage inequality as measured by the Gini coefficient fell in a statistically significant way between 2014 and 2018 (by -0.020, see column two). Note that this was the first decline in hourly wage inequality after a long period of substantial increase (Biewen and Seckler, 2019).

Further results in column two of table 2.2 show that, depending on the bite measure, the drop by -0.020 Gini points was more than fully explained by the minimum wage if one does not apply the trend adjustment (-0.035 for the regional, -0.027/-0.026 for the occupational/industry bite). As already suggested by the graphical analysis, the pre-trend correction results in lower minimum wage effects. Still, applying the two-year trend adjustment suggests that the minimum wage either fully or largely explained the drop in hourly wage inequality between 2014 and 2018 (-0.022 for the regional bite, -0.017 for the occupational/industry bite). These results suggest that, while the introduction of the minimum wage causally reduced wage inequality, the inequality trend from 2014 to 2018 would already have been flat without its introduction, implying that the minimum wage was not the only factor breaking the long-term trend of increasing wage inequality.

Table 2.2: Minimum wage effects on inequality in hourly wages, 2014 vs. 2018

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	16.247 (0.241)	0.260 (0.002)	3.504 (0.073)	2.015 (0.031)	1.739 (0.012)
2018	17.740 (0.208)	0.240 (0.002)	3.339 (0.080)	1.992 (0.030)	1.676 (0.018)
$\hat{\Delta}^{18-14}$	1.493*** (0.329)	−0.020*** (0.003)	−0.165 (0.113)	−0.023 (0.045)	−0.063*** (0.022)
Bite 1 (Regions)					
No trend adjustment	0.371*** (0.064)	−0.035*** (0.002)	−0.829*** (0.066)	−0.118*** (0.032)	−0.299*** (0.023)
1-year trend adjustment	0.309*** (0.066)	−0.029*** (0.002)	−0.670*** (0.063)	−0.116*** (0.032)	−0.225*** (0.023)
2-year trend adjustment	0.233*** (0.074)	−0.022*** (0.002)	−0.529*** (0.060)	−0.114*** (0.033)	−0.160*** (0.022)
Bite 2 (Augmented occupations)					
No trend adjustment	0.032 (0.071)	−0.027*** (0.006)	−0.750*** (0.109)	−0.050 (0.038)	−0.326*** (0.041)
1-year trend adjustment	0.053 (0.071)	−0.022*** (0.005)	−0.635*** (0.102)	−0.057 (0.038)	−0.263*** (0.037)
2-year trend adjustment	0.056 (0.112)	−0.017*** (0.005)	−0.534*** (0.106)	−0.062 (0.045)	−0.209*** (0.036)
Bite 3 (Augmented industries)					
No trend adjustment	0.075* (0.043)	−0.026*** (0.004)	−0.762*** (0.118)	−0.057 (0.042)	−0.325*** (0.052)
1-year trend adjustment	0.085* (0.047)	−0.021*** (0.003)	−0.638*** (0.119)	−0.064 (0.045)	−0.258*** (0.051)
2-year trend adjustment	0.088 (0.067)	−0.017*** (0.003)	−0.502*** (0.118)	−0.070 (0.053)	−0.186*** (0.051)

Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Note that this conclusion depends on the inequality measure chosen. For the Q90/Q10 and the Q50/Q10 ratio (columns three to five of table 2.2), the inequality reducing effect of the minimum wage exceeds the actual fall in inequality, suggesting that inequality as measured by these quantile gaps would have risen without the minimum wage. Columns three to five of table 2.2 also indicate that the introduction of the minimum wage specifically reduced inequality in the lower half of the hourly wages distribution as measured by the Q50/Q10 ratio. For the regional bite definition, there is also a significant effect for the upper half (as measured by the Q90/Q50 ratio) owing to the fact that measured spillovers are stronger and even reach beyond the median hourly wage (around 16 euros/hour in 2018). We conclude that the minimum wage introduction explains a very large share of the decline in the inequality of hourly wages, a share which is larger than the share explained in the reduction in monthly earnings as found by Bossler and Schank (2023).

2.5.2 Effects on hours worked

We consider three different subgroups reflecting potential differences in how working hours might react to the minimum wage introduction: i) workers with hourly wages below 12 euros/hour, ii) workers with hourly wages between 12 and 16 euros/hour, and iii) workers with hourly wages above 16 euros/hour. In the following, we will only discuss selected results for groups i) and ii). For group iii), we obtain sharply measured zero effects throughout, which we document in full in the appendix. Again, the minimum wage effects being zero for high-wage earners can be interpreted as a validity check for our estimation procedure.

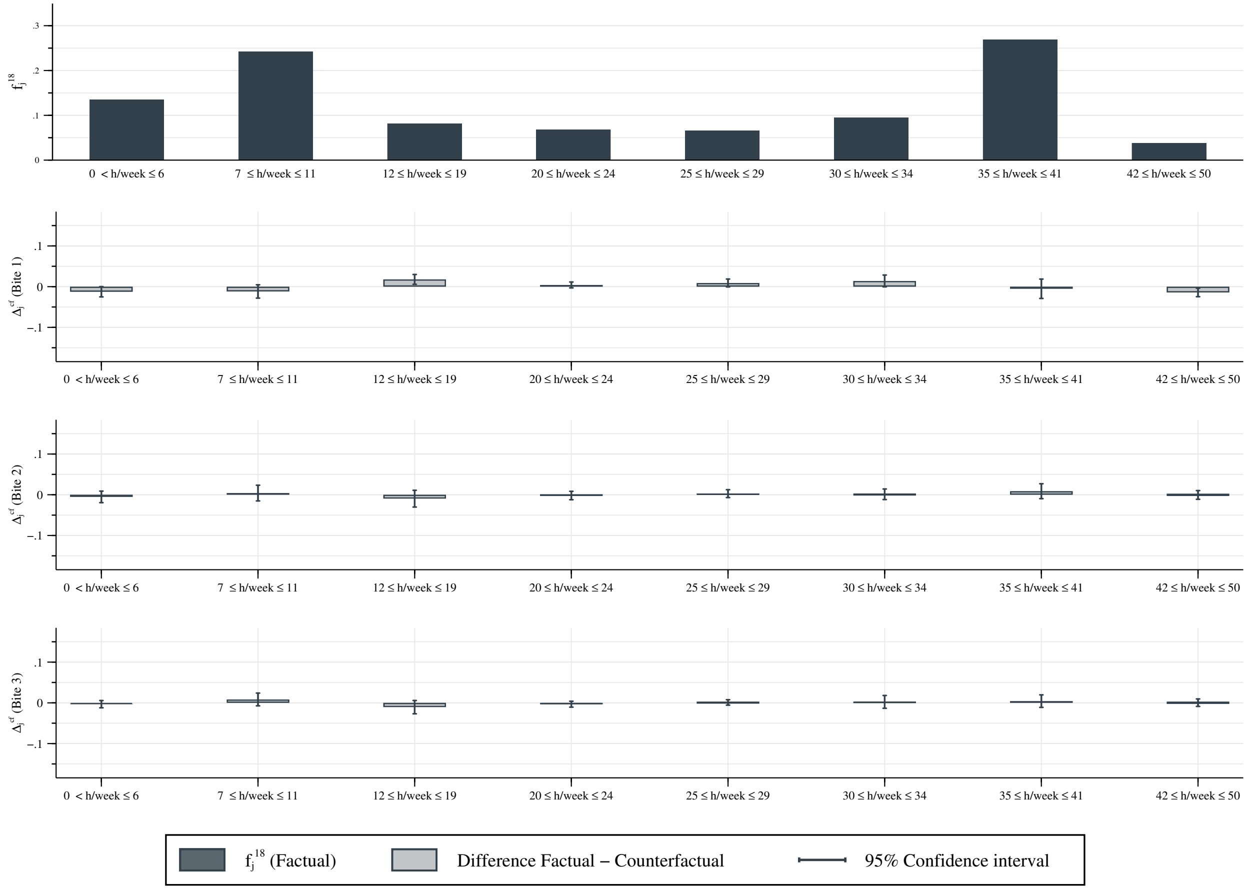
Figure 2.5 presents the results for the worker group with hourly wages below 12 euros/hour. For the regional bite definition, there are marginally significant positive effects for the fraction of hours worked in the interval 12 to 19 hours per week, and marginally significant negative effects for 42 to 50 hours per week (second panel of figure 2.5). By contrast, the effects are insignificant and close to zero in all other remaining cases for the regional bite and in all cases for the other two bite definitions used (third and fourth panel of figure 2.5). The corresponding results for the group with hourly wages between 12 and 16 euros/hour shown in figure 2.6 also show insignificant zero effects across all hours bins and bite definitions.

There are less pronounced pre-trends for hours worked than for hourly wages (see appendix figures A.20/A.35 as well as figures A.21/A.36). They are almost flat and largely insignificant for the occupational and industry bite. For the regional bite, there are some more pronounced but nonlinear and concave patterns. For wages up to 12 euros/hour, high-bite regions show an increase in the distribution function for the low- and medium-hours bins – the effect being strongest for the 7 to 11 hours bin, i.e., the share of workers with lower hours of work is increasing between 2011 and 2014 with a particular strong increase between 2011 and 2012. For wages between 12 and 16 euros/hour, high-bite regions show concave pre-trends with positive trends in various bins between 2011 and 2012 and falling pre-trends between 2012 and 2014 for medium-hours bins but not in the lowest hours bin. Due to the concavity of the pre-trends, the case for a linear trend extrapolation is much weaker for hours of work compared to hourly wages (and, as we see below, compared to monthly earnings). Still correcting for pre-trends regarding the regional bite does not change the significant effects in two bins for wages up to 12 euros/hour and the zero (insignificant) effects in all other bins (figures A.17 and A.32). To complete the discussion, correcting for pre-trends in the cases of the occupational bite and the industry bite confirms the previous finding of no effects of the minimum wage (lower panels of figures A.18/A.33 and A.19/A.34).

Our findings suggest that the bite definition makes a difference for hours of work. For the regional bite, the minimum wage introduction did increase the share of workers in the 12 to 19 hours bin and reduce the share of those working 42 hours and more among workers making at most 12 euros/hour. Thus, within a region, low-wage workers shifted away from low-paid jobs with long hours to jobs with lower hours that became comparatively better paid after the introduction of the minimum wage. However, this shift did not occur within industries and occupations, suggesting that the effect found for the regional bite is due to reallocation effects within regions but across occupations and industries (Dustmann et al., 2022).

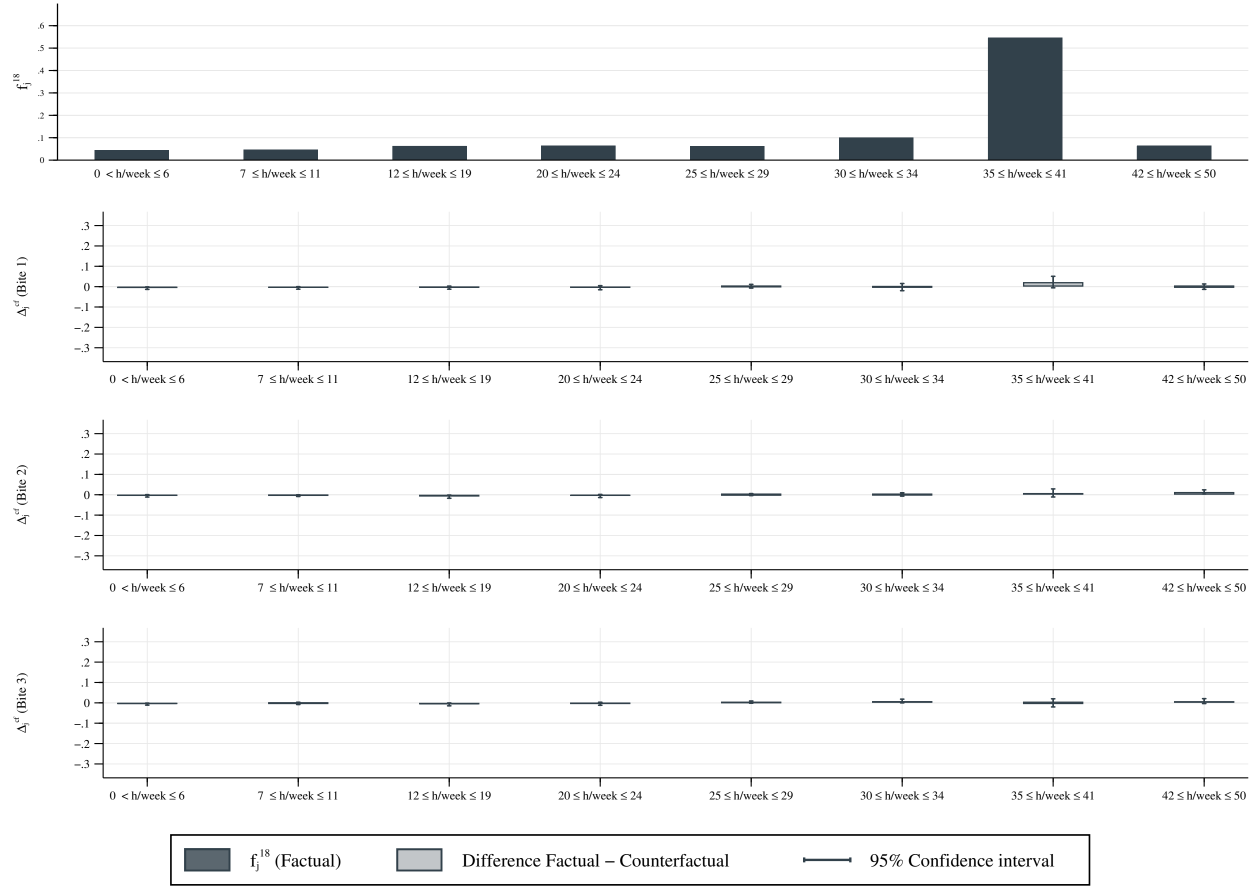
In sum, our results on the effects on hours worked are more mixed than those for hourly wages. We find no statistically significant effects of the minimum wage on the distribution of working hours based on the occupational and the industry bite. By contrast, there are significant effects for two specific hours bins (12 to 19 hours, 42 to 50 hours) for the regional bite. These suggest a reduction in working hours due to the minimum wage introduction. These findings suggest that the shift from full-time work with overtime hours to part-time work with 12 to 19 hours is caused by reallocation effects at the regional level across industries and occupations, not by hours reductions to keep wage bills constant.

Figure 2.5: 2018 Factual distribution of weekly working hours and treatment effect due to minimum wage for individuals with hourly wages ≤ 12 euros/hour.



Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show differences between the factual and counterfactual bin frequencies for different bite specifications. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure 2.6: 2018 Factual distribution of weekly working hours and treatment effect due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour.



Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show differences between the factual and counterfactual bin frequencies for different bite specifications. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

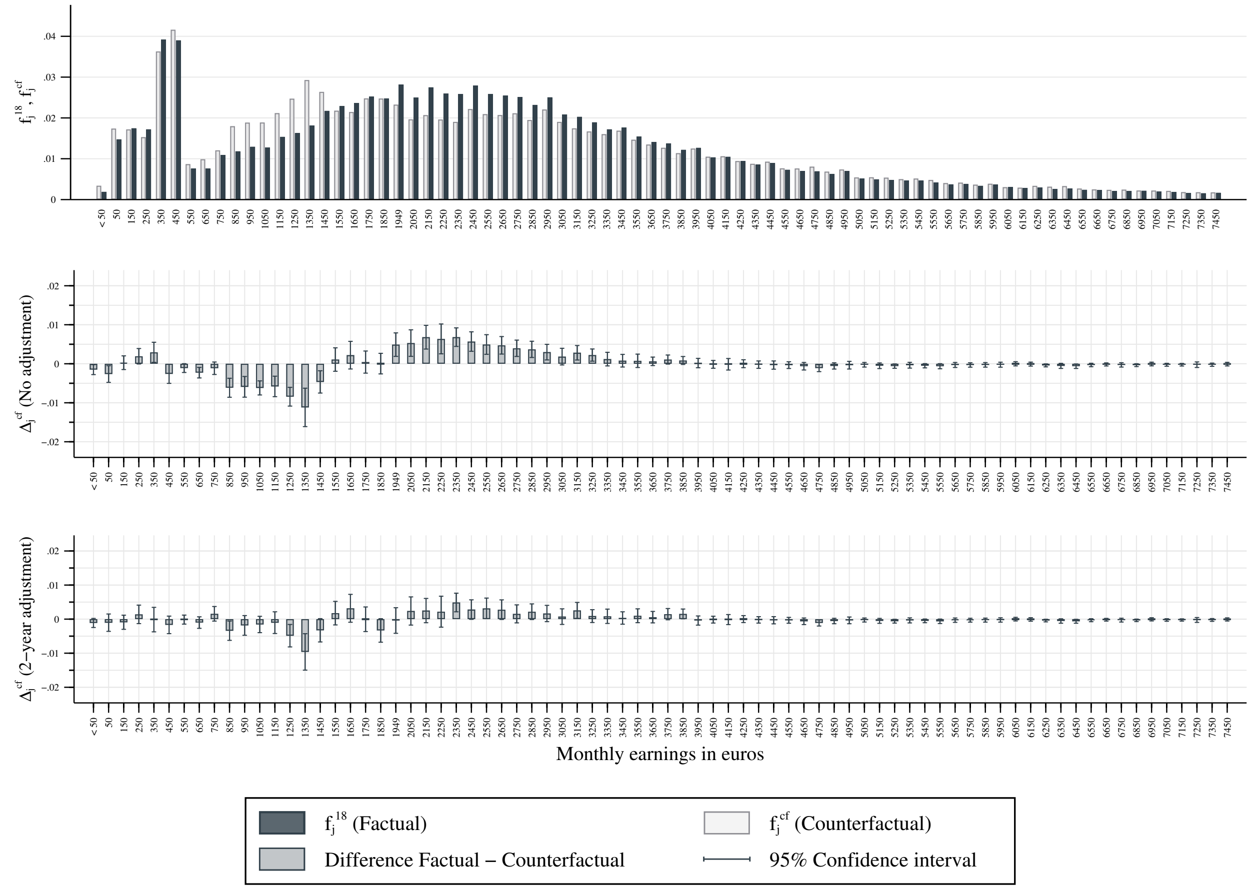
2.5.3 Effects on monthly earnings

In this section, we consider the effects of the minimum wage introduction on monthly earnings. These have also been studied based on administrative IAB data by Bossler and Schank (2023). The effects of the minimum wage on monthly earnings reflects the joint effect of changes in hourly wages and in hours worked. Given that we observe little evidence for systematic changes of working hours, we would expect significant changes in the distribution of monthly earnings. However, the question is how strong these changes where and which workers in the distribution benefitted most.

Our results for monthly earnings are shown in figures 2.7 to 2.9. The upper panels show factual and counterfactual monthly earnings distributions for 2018, while the two lower panels display the causal effects of the minimum wage introduction on the earnings distribution in terms of differences at each bin (with and without pre-trend correction). Figure 2.7 for the regional bite suggests that the introduction of the minimum wage benefitted workers with very low monthly earnings (up to 450 euros per month, marginal part-time) but to a limited extent. The main effect was for workers in the lower middle part of the distribution (850 to 1,450 euros per month), whose earnings shifted to levels around and above the median (2,472 euros per month in 2018, compare middle panel of figure 2.7). The pattern looks very similar when using the occupational and industry bite but the measured effects are noticeably weaker, again suggesting the strongest spillover effects at the regional level.

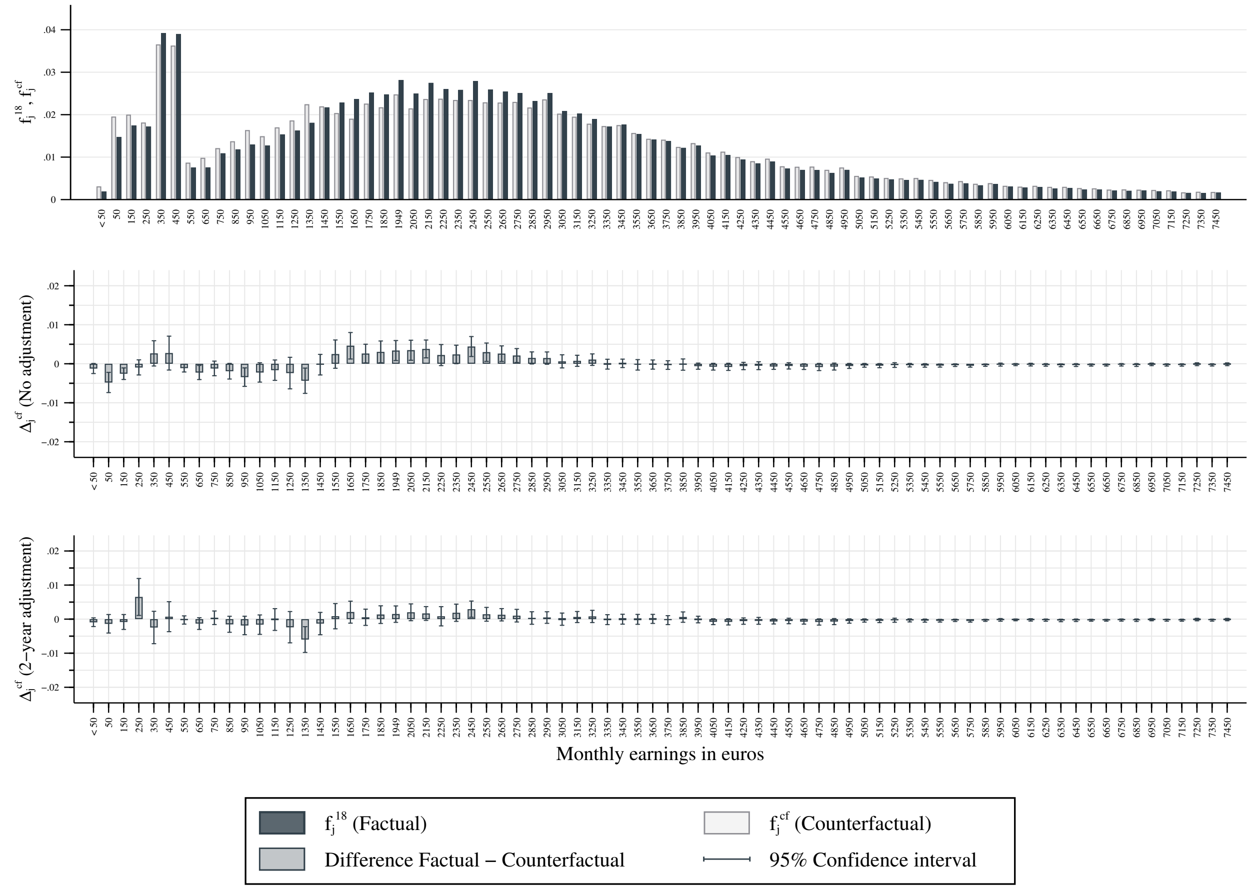
Compared to hourly wages, we find less clear and often nonlinear pre-trends suggesting higher earnings growth in high-bite groups for moderately low and medium earnings levels (see figures A.65 to A.71 in the appendix). Similar to hours of work, the pre-trends turn out concave for the regional bite. In contrast, they are convex for the two other bites. A linear pre-trend specification between 2012 and 2014 seems a good approximation. Accounting for these pre-trends significantly weakens the observed effects on the earnings distribution, see lower panels of figures 2.7 to 2.9.

Figure 2.7: 2018 Factual vs. counterfactual distribution of monthly earnings in the absence of minimum wage. Bite 1: Regions.



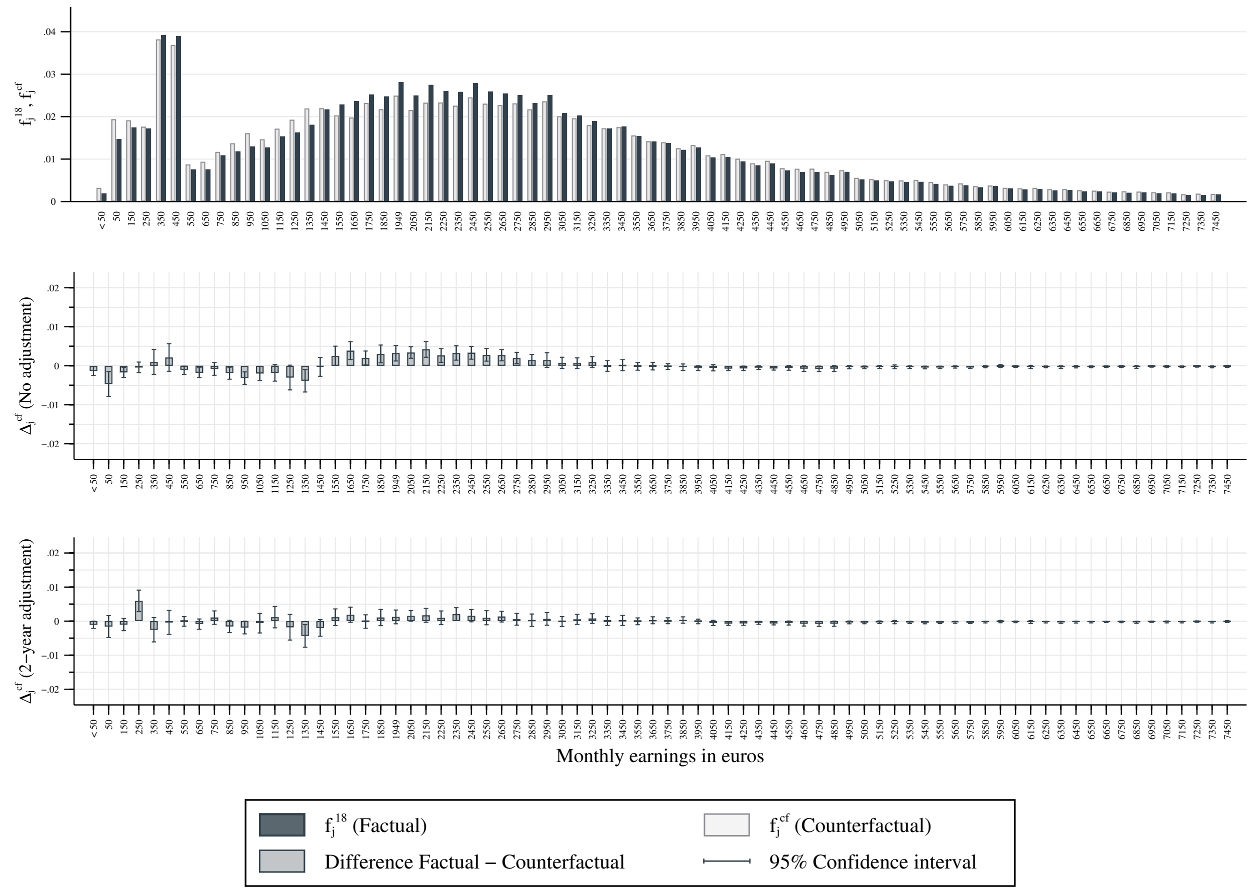
Notes: The x-axis shows monthly wage bins. For example, the ‘1050’ bin comprises monthly earnings in the interval [1,050; 1,149] euros. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure 2.8: 2018 Factual vs. counterfactual distribution of monthly earnings in the absence of minimum wage. Bite 2: Augmented occupations.



Notes: The x-axis shows monthly wage bins. For example, the ‘1050’ bin comprises monthly earnings in the interval [1,050; 1,149] euros. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure 2.9: 2018 Factual vs. counterfactual distribution of monthly earnings in the absence of minimum wage. Bite 3: Augmented industries.



Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The counterfactual bins in the first row of the figure correspond to the model-implied counterfactual distributional mass in the absence of the minimum wage *without trend adjustment*. The second and third panel show differences in bin frequencies (second panel: no trend adjustment, third panel: 2-year trend adjustment). 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Table 2.3 summarizes the effects of the minimum wage introduction on earnings inequality. Ignoring pre-trends, the fall in the Gini by -0.020 is fully explained by the minimum wage when using the regional bite (-0.020) but only half explained when using the occupational or industry bite (-0.012, see second column of table 2.3), see Bossler and Schank (2023) for similar findings.

Table 2.3: Minimum wage effects on inequality in monthly earnings, 2014 vs. 2018

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	2305.181 (32.207)	0.355 (0.002)	11.566 (0.342)	2.131 (0.032)	5.427 (0.117)
2018	2483.971 (27.681)	0.336 (0.002)	10.991 (0.302)	2.108 (0.031)	5.215 (0.101)
$\hat{\Delta}^{18-14}$	178.791*** (43.967)	−0.020*** (0.002)	−0.575 (0.465)	−0.024 (0.046)	−0.212 (0.157)
Bite 1 (Regions)					
No trend adjustment	56.948*** (9.801)	−0.020*** (0.002)	−0.519*** (0.137)	−0.194*** (0.035)	0.215*** (0.072)
1-year trend adjustment	48.948*** (10.268)	−0.016*** (0.002)	−0.615*** (0.165)	−0.179*** (0.032)	0.139* (0.078)
2-year trend adjustment	39.683*** (10.088)	−0.012*** (0.002)	−0.676*** (0.211)	−0.165*** (0.030)	0.081 (0.095)
Bite 2 (Augmented occupations)					
No trend adjustment	−2.521 (14.507)	−0.012*** (0.003)	−0.906*** (0.334)	−0.076 (0.050)	−0.233* (0.130)
1-year trend adjustment	−0.410 (14.976)	−0.009*** (0.003)	−0.644** (0.256)	−0.087* (0.050)	−0.088 (0.102)
2-year trend adjustment	−1.768 (19.378)	−0.006** (0.003)	−0.409 (0.296)	−0.096* (0.050)	0.041 (0.127)
Bite 3 (Augmented industries)					
No trend adjustment	2.222 (8.641)	−0.012*** (0.002)	−0.917*** (0.233)	−0.075* (0.043)	−0.242*** (0.083)
1-year trend adjustment	2.835 (8.602)	−0.009*** (0.002)	−0.695*** (0.205)	−0.083** (0.041)	−0.121* (0.072)
2-year trend adjustment	2.920 (12.785)	−0.006*** (0.002)	−0.492** (0.205)	−0.089** (0.040)	−0.012 (0.078)

Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

After pre-trend correction, these contributions shrink to -0.012 for the regional bite and to -0.006 for the occupational and the industry bite. Columns three and four of table 2.3 show that the minimum wage significantly reduced both the Q90/Q10 and the Q90/Q50 ratio as both median and very low earnings increased. The last columns of table 2.3, suggest an *increase* of the Q50/Q10 ratio when using the regional bite, but a *decrease* based on the other two bite definitions. The explanation for this finding is that the stronger spillovers in the case of the regional bite led to a more pronounced increase in median earnings compared to the two other bite definitions. The opposing effects lose statistical significance when correcting for pre-trends.

Taken together, our results for monthly earnings concerning the regional bite are very much in line with those in Bossler and Schank (2023), who also find some but limited effects at the bottom of the distribution, no effects in the range between the bottom and the lower middle of the distribution, and the strongest effects of the minimum wage in the lower middle of the earnings distribution up to the median. However, compared to the strong minimum wage effects found by us on hourly wages, the effects on monthly earnings are somewhat muted, especially after the adjustment for pre-trends and when considering the occupation and industry bite definitions. This may be due to the fact that minimum wage earners are a smaller share of workers at each point of the overall distribution of monthly earning. Thus, the impact of the minimum wage introduction is weaker at each point of the overall distribution and more dispersed along the distribution in contrast to the distribution of hourly wages where the direct impact can be pinned down exactly at a certain value of the hourly wage where one expects a spike in the post minimum wage distribution. This could also explain why Caliendo et al. (2023)'s study based on survey data with relatively small sample sizes did not find significant effects on monthly earnings.

2.5.4 Sensitivity analysis

Given the possibility that the minimum wage potentially also changed worker characteristics on which we condition in our analysis, we investigate the sensitivity of our results with respect to varying sets of conditioning variables (see appendix for full set of results). A point in case are re-allocation effects between firms as studied by Dustmann et al. (2022). In a first sensitivity analysis, we therefore omitted all firm variables from our conditioning set. This led to almost identical estimates in all scenarios. We therefore conclude that our distributional results are robust with respect to re-allocation effects between firms alone. In a second sensitivity analysis, we omitted *all* conditioning variables from our distribution regressions (i.e., we included only the time and group effects as well as the interactive difference-in-differences terms). This led to small changes in estimated effects, but left our qualitative conclusions unchanged. Omitting conditioning variables led to somewhat more pronounced inequality reducing effects of the minimum wage on the distributions of hourly wages (compare table 2.2 to table A.7 in the appendix) and monthly earnings (compare table 2.3 to table A.10 in the appendix). Another natural consequence of not conditioning on covariates was the noticeably lower precision of the point estimates (conditioning reduces the error variance). Note that conditioning on covariates also eliminates effects of compositional changes in workforce characteristics on the outcome distribution. Given that there were only few compositional changes between 2014 and 2018 (see table A.2 in the appendix), it is not surprising that results change only little when not conditioning on characteristics.

2.6 Discussion and conclusion

This paper analyzes the effects of the German statutory minimum wage on the distributions of hourly wages, hours worked, and monthly earnings. Our analysis is based on the German Structure of Earnings Survey (GSES) and administrative DGUV-IAB data, which are the only large-scale databases for Germany including information on hourly wages and working hours both before and after the minimum wage introduction. As a transparent methodological approach, we suggest to use difference-in-differences distribution regressions (DR-DiD) based on different bite measures for a full distributional analysis of minimum wage effects, while accounting for discrete mass points and changing nominal target values in these distributions. Using different bite measures allows to investigate the sensitivity of the estimated minimum wage effects based on the treatment intensity as measured in different segments of the labor market and to explore different channels for spillover effects.

Our results imply that the introduction of the minimum wage in 2015 caused low hourly wages to rise above its value and that it resulted in significant spillover effects up to 50-70 percent above. Given that we consider our information on wages and hours to be much less prone to rounding and other measurement error than in small-scale survey data with non-compulsory participation, our analysis indicates that such spillover effects are real. We find that wage inequality fell between 2014 and 2018, counteracting a long-term trend until 2010 in rising inequality in hourly pay and monthly earnings (Antonczyk et al., 2010; Biewen and Seckler, 2019; Bossler and Schank, 2023). Our results suggest that the introduction of the minimum wage explains a large part of the fall in inequality, depending on the inequality measure used. However, inequality as measured by the Gini would not have increased between 2014 and 2018 in the absence of the minimum wage. For the lower part of the distribution, we demonstrate that wage growth was already higher in groups that were later most affected by the minimum wage. The latter effect leads to an overestimation of minimum wage effects if these pretrends are not corrected for. All this may suggest that the minimum wage was not the only factor stopping the long-term trend of rising wage inequality. Such an interpretation is consistent with evidence in Biewen and Seckler (2019) who showed that de-unionization and compositional changes with respect to educational qualifications and work experience were responsible for rising inequality before 2014, but that the effect of these inequality drivers flattened out in the years before 2014.

Our comparison of alternative bite measures suggests the existence of substantial regional spillovers in wage determination as distributional effects on hourly wages and monthly earnings follow exactly the same patterns when a bite definition based on occupation or industry is used but lead to noticeably larger distributional impacts when the bite is defined at the regional level. Our results for working hours are more mixed in that we observe no working hours effects for the occupational and the industry bite but signficant effects for individual hours bins when using the regional bite, suggesting a slight shift from full-time work to part-time work which would be consistent with reallocation effects at the regional level (Dustmann et al., 2022).

Our overall conclusion is that the minimum wage changed prices (= hourly wages) while having only a small effects on quantities (= hours worked) so that the hourly wage increases – which we show to have affected hourly wages substantially above the minimum wage – also changed monthly earnings of workers. We show that these effects were strongest for workers not located at the very bottom, but in the lower middle and the middle of the earnings distribution, ranging to levels substantially above the median. However, effects on the distribution of monthly earnings also look weaker than the direct effects on the distribution of hourly wages, especially after adjustment for pre-trends. Our results thus help reconcile conflicting findings in the previous literature based on German administrative data (Bossler and Schank, 2023, who find significant effects of the minimum wage on monthly earnings), and German survey data (Caliendo et al., 2023, who do not find significant effects on monthly earnings, arguing that hourly wage increases did not translate into changes in monthly earnings because working hours were adjusted downwards). There is concern that non-compliance with the minimum wage may have led firms to reduce paid hours of work and not actual hours worked while keeping monthly earnings constant (Burauel et al., 2019). Our data involve paid hours and our findings of no hours of work effects for the industry and occupation bite speak against this concern because one would have expected a shift towards lower hours of works, especially in the low-hours bins.

Appendix

A.1 Identification assumptions for DR-DiD

In this section, we show that, when conceptualizing the distributional treatment effect problem as a distribution regression difference-in-differences model, identification is implied by the standard difference-in-differences assumptions for repeated cross-sections. Let I^z denote the dummy variable indicating whether or not the observed outcome Y is less than or equal to the threshold z, i.e., $\mathrm{I}^z = \mathbb{1}[Y \leq z]$. The potential outcome under treatment level $Bite = b$ is defined as $Y(b)$, and, correspondingly, $\mathrm{I}^z(b) = \mathbb{1}[Y(b) \leq z]$. Recall that there are two time periods $t = 0$ and $t = 1$ represented by the indicator $D_t = 0$ (for $t = 0$) and $D_t = 1$ (for $t = 1$). We assume repeated cross-section sampling, i.e., we observe iid samples from $(\mathrm{I}^z, Bite, W) \,|\, D_t = 0$ and from $(\mathrm{I}^z, Bite, W) \,|\, D_t = 1$, where W includes individual characteristics and time effects.

Recall that the factual distribution of Y in $D_t = 1$ is given by

$$F(z \,|\, D_t = 1) = \int \mathbb{E}(\mathrm{I}^z(b) \,|\, Bite = b, W, D_t = 1)\, dF(Bite, W \,|\, D_t = 1). \tag{A.1}$$

The counterfactual distribution under the assumption of no minimum wage is defined as

$$\begin{aligned} F^{cf}(z \,|\, D_t = 1) = & \int \mathbb{E}(\mathrm{I}^z(0) \,|\, Bite = b, W, D_t = 1)\, dF(Bite, W \,|\, D_t = 1) \qquad \text{(A.2)} \\ = & \int \{ \mathbb{E}(\mathrm{I}^z(b) \,|\, Bite = b, W, D_t = 1) \\ & - \underbrace{[\mathbb{E}(\mathrm{I}^z(b) - \mathrm{I}^z(0) \,|\, Bite = b, W, D_t = 1)]}_{:= ATT^z(b \,|\, b, W)} \}\, dF(Bite, W \,|\, D_t = 1) \end{aligned}$$

The parameter $ATT^z(b \,|\, b, W)$ is the average treatment effect for $Bite = b$ vs. $Bite = 0$ for individuals with characteristics W who actually receive treatment b, see Callaway and Sant'Anna (2021). Note that our research question involves only the comparison between treatment level $Bite = b$ and treatment level $Bite = 0$, so that the complications due to comparing different treatment levels (with nonzero bites) discussed in Callaway and Sant'Anna (2021) do not arise.

The following arguments identify $ATT^z(b\,|\,b,W)$, analogous to Callaway and Sant'Anna (2021):

$$\begin{aligned} ATT^z(b\,|\,b,W) &= \mathbb{E}(\mathrm{I}^z(b) - \mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 1) \\ &= \mathbb{E}(\mathrm{I}^z(b)\,|\,Bite = b, W, D_t = 1) - \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 1) \\ &= \mathbb{E}(\mathrm{I}^z(b)\,|\,Bite = b, W, D_t = 1) - \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 0) \\ &\quad - [\mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 1) - \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 0)] \end{aligned}$$

The last line in the above expression can not be estimated directly from the data. In addition to common support conditions and a no anticipation assumption in $E(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 0)$ (individuals who would be treated in $t = 1$ show outcome $\mathrm{I}^z(0)$ in $t = 0$), the key assumption used to identify the last line above is

$$\begin{aligned} &\mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 1) - \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = b, W, D_t = 0) \\ &= \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = 0, W, D_t = 1) - \mathbb{E}(\mathrm{I}^z(0)\,|\,Bite = 0, W, D_t = 0), \end{aligned} \tag{A.3}$$

i.e., in the treated group, wage growth at different points of the distribution in the absence of treatment would be the same as in the untreated group. Replacing the last line for $ATT^z(b\,|\,b,W)$ by (A.3), allows to estimate $ATT^z(b\,|\,b,W)$.

Our motivation for assumption (A.3) in our application is as follows. Take the case in which the intensity of treatment is defined by the minimum wage bite at the regional level. In this case, W contains productivity characteristics such as education, experience, occupation, industry etc. Then (A.3) amounts to assuming that wage changes for workers in narrow education/experience/occupation/industry etc. cells evolve in a parallel fashion across different regions in the absence of a minimum wage. If systematic deviations from this assumption are observed in pre-treatment periods, then the extrapolations of such a trend be incorporated into the above expressions (this is what we do in section 2.4.2, analogous to Dobkin et al. (2018); Ahlfeldt et al. (2018); Freyaldenhoven et al. (2021) for the non-distributional case).

Condition (A.3) is the conditional version of the condition identified by Roth and Sant'Anna (2023) to characterize the situation that parallel trends are insensitive to functional form (i.e., to strictly monotonic transformations of the outcome). This condition is a 'parallel trends-type assumption for the cumulative distribution function of untreated potential outcomes' and is stated in Roth and Sant'Anna (2023) for the case of two treatment levels and no covariates as $F_{Y_1(0)\,|\,treatment=1}(y) - F_{Y_0(0)\,|\,treatment=1}(y) = F_{Y_1(0)\,|\,treatment=0}(y) - F_{Y_0(0)\,|\,treatment=0}(y)$ (proposition 3.1 in Roth and Sant'Anna, 2023). To see the equivalence to (A.3), recall that cumulative distribution functions of Y are defined as $F(z\,|\,\cdot) = \mathbb{E}(\mathrm{I}^z\,|\,\cdot)$. This type of identification condition represents a substantial improvement over earlier approaches to find identification assumptions for distributional treatment effects in that it avoids restrictions on the joint distribution of outcomes in $t = 0$ and $t = 1$ (e.g., Callaway and Li, 2019; Fan and Yu, 2012). Hence, it easily extends to the cross-sectional case. Note the implication that DR-DiD is automatically invariant to functional form of the outcome, which directly follows from the fact that threshold indicators are unchanged by monotonic transformations, e.g., $\mathbb{1}[y \leq z] = \mathbb{1}[\log(y) \leq \log(z)] = \mathbb{1}[y^* \leq z^*]$.

The original contribution by Roth and Sant'Anna (2023) provides an interpretation of the 'parallel trends assumption for the cumulative distribution function of untreated potential outcomes', implying that the underlying data generating process involves a mixture of random assignment and stationary potential outcomes. This strong interpretation has been questioned by Kim and Wooldridge (2024) who show that the condition only requires that distributional change in the treatment and control group has a common component absent treatment. Note that we only use a conditional version of the original condition which is substantially weaker irrespective of its exact interpretation.

Note that we impose in our actual application the additional assumption $ATT^z(b\,|\,b,W) = ATT^z(b\,|\,b) = \beta_z \cdot Bite$. This entails two substantial restrictions, which we impose for practical and statistical reasons. The first restriction is that the treatment effect is independent of W (homogeneity). In principle, this could be relaxed, but we found this to be difficult both practically and statistically given the many covariates in W. Relaxing this restriction would also substantially complicate the pre-trend analysis (which would have to be carried out separately by subgroups characterized by W). The second restriction is that the treatment effect is linear in treatment intensity. In principle, this could be relaxed by discretizing treatment intensity. However, when experimenting with different ways to do this, we found that discretizing the bite variable into a non-trivial number of categories quickly introduces a lot of noise into the estimations. It also complicates the pre-trend analysis considerably. Unfortunately, given the computational limitations we face due to the restricted on-site access to our databases, we have to abstain from pursuing more flexible approaches in our application. In line with Roth and Sant'Anna (2023), we also point out that, despite its potential limitations, the linear DiD specification is still by far the most widely used model DiD design with continuous treatment variables.

A.2 Testing the shape restriction on the counterfactual distribution

As in Roth and Sant'Anna (2023) and Kim and Wooldridge (2024), our setup has the testable implication that the resulting counterfactual cumulative distributions (2.4) and (2.12) are proper distribution functions. Following Kim and Wooldridge (2024), we subject this implication to the formal test developed by Chen and Szroeter (2014). In our case, the test involves assessing the null hypothesis that all J counterfactual grouped probability functions defined in (2.6) are simultaneously greater than or equal to zero, against the alternative hypothesis that at least one of the bins has an implied negative distributional mass. Formally, this is given by

$$\mathrm{H}_0 : f^{cf}_{1,1} \geq 0, f^{cf}_{2,1} \geq 0, \ldots, f^{cf}_{J,1} \geq 0. \quad \mathrm{H}_1 : f^{cf}_{j,1} < 0, \text{ for at least one } j = 1, \ldots, J. \tag{A.4}$$

As outlined in Chen and Szroeter (2014) and Kim and Wooldridge (2024), the procedure involves a straightforward test for multiple inequality hypotheses, avoiding complex calculations since indicator functions are smoothed. As formally shown in Chen and Szroeter (2014), under the null hypothesis, this statistic asymptotically follows a normal distribution leading to a standard test procedure.

In the following, we describe the procedure for testing the null hypothesis in (A.4). Let $\widehat{V}$ be the estimated $(J \times J)$ variance-covariance matrix for all grouped probability bins $f^{cf}_{1,1}, \ldots, f^{cf}_{J,1}$. Denote the estimated variances on the main diagonal as $\widehat{v}_{jj}$. In practice, this estimate is obtained using the clustered bootstrap approach described above, with clusters referring to the respective level of treatment. Denote the number of clusters as M. Furthermore, following the notation in Chen and Szroeter (2014), the following terms are defined:

$$\widehat{\theta}_j \equiv \frac{1}{\sqrt{\widehat{v}_{jj}}}, \quad \widehat{\Delta} \equiv \text{diag}(\widehat{\theta}_1, \ldots, \widehat{\theta}_J).$$

As described in Chen and Szroeter (2014), earlier contributions proposed test procedures that involve using indicator functions to handle the inequality constraints implied by the null hypothesis, leading to non-standard distributions of the test statistic. To overcome this, Chen and Szroeter (2014) propose a smoothing approach that approximates the indicator but ultimately ensures the asymptotic normality of the test statistic. Technically, Chen and Szroeter (2014) consider a left-tailed test that is built on the quantity

$$\sum_{j=1}^{J} \left[\sqrt{M} \Psi_M(\widehat{f}_{j,1}^{cf}) \widehat{f}_{j,1}^{cf} - \Lambda_M(\widehat{f}_{j,1}^{cf}, \widehat{v}_{jj}) \right],$$

with $\Psi_M(x) \equiv \Psi(K(M)x)$ being the smoothed indicator. For later reference, denote $\psi(x)$ to be the first derivative of $\Psi(x)$. $K(M)$ is a tuning parameter that grows with the sample size M. The conditions outlined in Chen and Szroeter (2014), ensure that the test statistic is either degenerate or approaches a normal distribution, eliminating the need for simulation. $\Lambda_M(\widehat{f}_{j,1}^{cf}, \widehat{v}_{jj})$ converges to zero but adjusts for potential size distortion in finite samples.

Following Chen and Szroeter (2014) and Kim and Wooldridge (2024), we use the following terms to define the test statistics. The smoothed indicator is defined to be logistic, i.e., $\Psi(x) \equiv (1+\exp(x))^{-1}$. Accordingly, $\psi(x) = -\exp(x)/((1+\exp(x))^2)$. The tuning function is given as $K(M) = \sqrt{M/\log(M)}$. The adjustment factor is defined as, $\Lambda_M(\widehat{f}_{j,1}^{cf}, \widehat{v}_{jj}) \equiv \widehat{v}_{jj} \psi\left(K(M)\widehat{f}_{j,1}^{cf}\right) K(M)/\sqrt{M}$. The final testing decision is based on

$$Q = \begin{cases} \Phi(Q_1/Q_2) & \text{if } Q_2 > 0 \\ 0 & \text{if } Q_2 = 0, \end{cases} \tag{A.5}$$

$$\text{with} \quad Q_1 \equiv \sqrt{M}\widehat{\Psi}'\widehat{\Delta}\widehat{f}_1^{cf} - e_J'\widehat{\Lambda},$$

$$\text{and} \quad Q_2 \equiv \sqrt{\widehat{\Psi}'\widehat{\Delta}\widehat{V}\widehat{\Delta}\widehat{\Psi}},$$

where

$$\begin{aligned}
\widehat{\Psi} &\equiv \left(\Psi\left(K(M)\widehat{\theta}_1\widehat{f}^{cf}_{1,1}\right),\ldots,\Psi\left(K(M)\widehat{\theta}_J\widehat{f}^{cf}_{J,1}\right)\right)',\\
\widehat{\Lambda} &\equiv \left(\Lambda_M(\widehat{f}^{cf}_{1,1},\widehat{v}_{11}),\ldots,\Lambda_M(\widehat{f}^{cf}_{J,1},\widehat{v}_{JJ})\right)',\\
\widehat{f}^{cf}_1 &\equiv \left(\widehat{f}^{cf}_{1,1},\ldots,\widehat{f}^{cf}_{J,1}\right)',\\
e_J &\equiv (1,\ldots,1)',
\end{aligned}$$

and $\Phi(\cdot)$ is the cdf of the standard normal distribution. Under the null, Q_2 is generally positive, so the ratio Q_1/Q_2 is asymptotically normal. The outcome of the test statistic can be interpreted as a p-value.

In our application, we generally do not reject the null hypothesis that our counterfactual distributions are proper distribution functions at any conventional significance level for all the many specifications considered. An exception is the specific case of hourly wages with two-year trend-adjustment. Closer inspection shows that, in this case, probability masses for the two lowest hourly wages bins $[0;3.49]$ and $[3.50;4.49]$ turn slightly negative in a statistically significant way. Given the very low initial probabilities for these wage bins, this indicates 'trend over-correction' for these cases at the very low end of the distribution (a positive $\bar{\Delta}_z$ gets subtracted from a value already close to zero in (2.12)). While we do not think that this observed irregularity at the very extreme end of the distribution (and only in connection with trend-adjustment) indicates a substantive problem in our application, the result warns about possible problems in more general applications as indicated in Fernández-Val et al. (2024).

A.3 Differences between DR-DiD and RIF-DiD

For the following, also see the discussion in Dube (2019b). The main differences between RIF-DiD and DR-DiD that lead us to adopt the DR-DiD approach in our application is that DR-DiD can deal with discrete mass points and nominal values of the outcome variable, while the RIF approach is based on continuous operations on continuous distributions which rule out these cases. Moreover, the RIF approach targets aggregate statistics such as quantiles and inequality measures rather than nominal levels of the outcome variable. If one is interested in particular nominal points of the outcome distribution then one could in principle define the quantiles that correspond to these points. However, this is not possible in a DiD setup as there are multiple time periods (e.g., quantiles will correspond to varying nominal points in the distribution in different time periods). Moreover, modeling quantiles in order to target nominal points would involve unnecessary inversions (from nominal points to quantiles and back) which are not nececssary in the DR approach.

Apart from these aspects, we list the following points to highlight the differences between DR and RIF when applied to a DiD setup. In general, recall that the recentered influence function of a statistic θ is defined as $RIF(y,\theta) = IF(y,\theta) + \theta$, where $IF(y,\theta)$ is the influence function of the statistic θ (Firpo et al., 2009). A difficulty of the RIF approach in the context of DiD is that the RIF regression involves different time periods (pre- and post-treatment) raising the question whether the estimate of θ used to recenter the influence function $IF(y,\theta)$ shall be computed only from the pre-treatment period or from the pooled sample (the latter potentially being affected by the treatment effects). Given that θ is typically a highly aggregated statistic, the difference between the two cases is probably small in many applications, but it may be large if the treatment has a big effect (as in our application). By contrast to the dependent variable of the RIF regresssion, the dependent variables of the DiD regression (i.e., the threshold indicators) do not use any distributional information but only information of the observation itself.

In terms of identification, we showed above that, in order to identify the full distributional treatment effect, the DR-DiD approach needs to make a parallel-trends assumption at each threshold indicator of the outcome distribution. The latter is known to be equivalent to the parallel-trends assumption being independent of the functional form of the outcome variable Y and the requirement of (as-if) treatment randomization and stationarity (Roth and Sant'Anna, 2023). One might wonder if the RIF-DiD approach is less restrictive as it only needs to invoke a parallel-trends assumption for the chosen form of the RIF-function. However, to recover the full distributional treatment effect, one has to compute the RIF-regression separately for a comprehensive set of quantiles (each quantile having its separate RIF-function). For this, one has to invoke a separate parallel-trends assumption for each quantile, which is equivalent to assuming parallel-trends assumptions for the set of thresholds that correspond to these quantiles. As a consequence, the RIF-DiD approach is as demanding as the DR-DiD approach if the goal is to identify the full distributional treatment effect.

If one is only interested in the treatment effect on a particular functional θ of the counterfactual distribution (and ignoring the problem of determining θ on the pooled sample), then RIF-DiD indeed only requires to make a parallel-trends assumption for the RIF-function of the statistic of interest $RIF(y,\theta)$ (e.g., the RIF of the Gini). This shortcut is not possible in the DR-DiD approach in which one always first has to identify the full distributional treatment effect and then possibly derives results for functionals from that. On the other hand, using this shortcut in the RIF-DiD approach will have to make the assumption that the usually highly nonlinear object $\theta = \mathbb{E}(RIF(Y,\theta))$ can be well-approximated by a DiD regression model. Depending on the application, this may be more restrictive than assuming that regression models for threshold indicators (or influence functions for individual quantiles) follow a DiD regression structure which then identify the full distributional treatment effect from which results for particular functionals can be derived (as we do in our empirical application).

A.4 Additional tables and figures

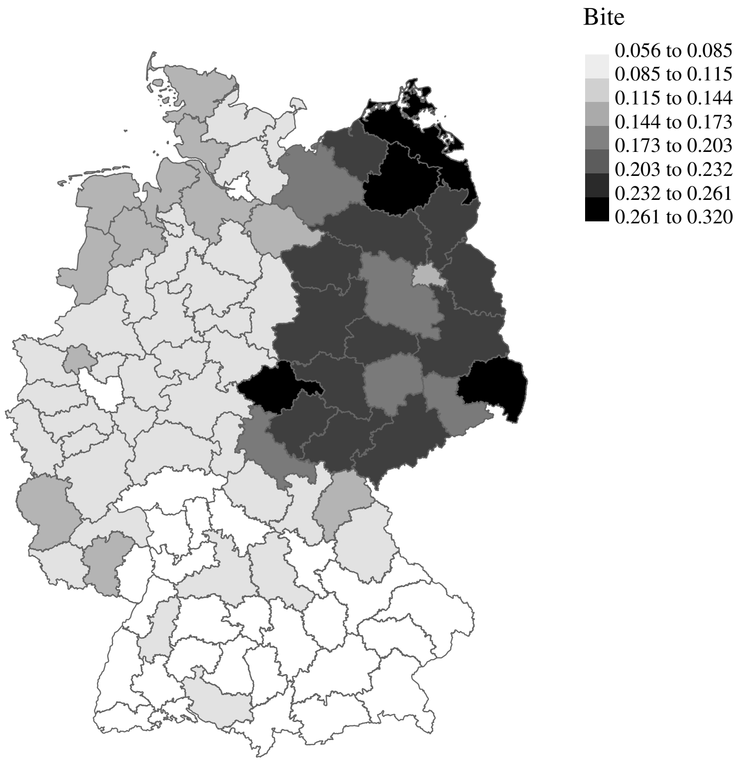
A.4.1 Highest and lowest bite categories

Table A.1: Bite descriptive statistics – Lowest and highest categories

	(1) German regions	(2) Occupations+East/West	(3) Industry+East/West
Lowest categories			
Lowest	München	Technical research, development, construction and production control occupations – West	Financial service activities + (Re)Insurance/pension funding – East
2nd lowest	Ingolstadt	Computer science, information and communication technology occupations – West	Manufacturing of motor vehicles, trailers and semi-trailers – West
3rd lowest	Oberland (Bavaria)	Professions in financial services, accounting and tax consulting – West	Manufacturing of basic metals – West
4th lowest	Nürnberg	Professions in geology, geography and environmental protection – West	Manufacturing of basic pharmaceutical products and pharmaceutical preparations – West
5th lowest	Südostoberbayern	Construction planning, architecture and surveying professions – West	Financial service activities + (Re)Insurance/pension funding – West
Highest categories			
Highest	Vorpommern	Tourism, hotel and catering occupations – East	Food and beverage service activities – East
2nd highest	Mecklenburgische Seenplatte	Cleaning professions – East	Gambling and betting activities – East
3rd highest	Oberlausitz-Niederschlesien	Food manufacturing and processing – East	Other personal service activities – East
4th highest	Nordthüringen	Horticultural professions and floristry – East	Accommodation – East
5th highest	Mittleres Mecklenburg/Rostock	Protection, security and surveillance occupations – East	Security and investigation activities – East

Source: GSES 2014.

Figure A.1: Bite of the minimum wage across German regions



Notes: Graph shows the fraction of individuals with hourly wages less than the 2018 minimum wage (8.84 euros/hour) in the pre-policy period (April, 2014) across German regions (*'Raumordnungsregionen'*) (dark = higher bite).
Source: GSES 2014, own calculations.

A.5 Sample selection criteria

We apply the following sample selection restrictions to both the GSES and the DGUV-IAB data. As to data-driven restrictions, we have to drop most individuals working in the public sector (education, public administration) as these cases lack in the GSES the precise regional information needed to assign the minimum wage bite when computed at the regional level. Furthermore, we dropped groups of individuals not eligible to the minimum wage (§22 MiLoG, Mindestlohnkommission, 2020). This includes individuals under the age of 18, apprentices, individuals completing a voluntary internship in the course of their formal education up to the length of 3 months, the long-term unemployed, and volunteer workers. Except for the long-term unemployed, all of these group can be identified in the GSES data. In addition, we omitted individuals over the age of 65 as well as homeworkers. For plausibility reasons, we further dropped individuals with hourly wages larger than 100 euros/hour if we had information that these individuals were associated with low skills (the marginally employed, unskilled employees with simple, schematic tasks, or semi-skilled employees with predominantly simple tasks). We also discarded individuals who were reported to earn an hourly wage smaller than 3 or larger than 1,000 euros/hour, as well as individuals who were reported to work more than 350 hours per month.

A.6 Descriptive statistics

Table A.2: Descriptive statistics (GSES-sample)

	2014		2018	
Variable	mean	sd	mean	sd
Male				
Yes	0.525	0.499	0.531	0.499
Age				
Age 18-25	0.081	0.273	0.079	0.27
Age 26-30	0.106	0.308	0.108	0.311
Age 31-35	0.109	0.311	0.114	0.317
Age 36-40	0.106	0.308	0.113	0.317
Age 41-45	0.128	0.334	0.109	0.312
Age 46-50	0.162	0.368	0.136	0.343
Age 51-55	0.147	0.354	0.154	0.361
Age 56-60	0.109	0.312	0.125	0.331
Age 61-65	0.052	0.222	0.062	0.241
Educational attainment				
No degree, with or w/o voc. training	0.029	0.169	0.030	0.172
Lower or middle secondary, w/o voc. training	0.080	0.272	0.077	0.266
Lower or middle secondary, with voc. training	0.605	0.489	0.584	0.493
Upper secondary (Abitur), w/o voc. training	0.028	0.165	0.029	0.169
Upper secondary (Abitur), with voc. training	0.126	0.331	0.133	0.339
Diploma/Master degree, PhD	0.132	0.338	0.147	0.354
Tenure with current firm				
Tenure $\leq$ 5 yrs	0.501	0.500	0.521	0.500
Tenure 6-10 yrs	0.170	0.376	0.170	0.376
Tenure 11-15 yrs	0.116	0.320	0.094	0.291
Tenure 16-20 yrs	0.072	0.258	0.081	0.272
Tenure 21-25 yrs	0.063	0.243	0.048	0.214
Tenure $>$ 25 yrs	0.078	0.268	0.088	0.283
Federal State				
Schleswig-Holstein	0.030	0.170	0.030	0.170
Hamburg	0.028	0.165	0.028	0.166
Lower Saxony	0.091	0.288	0.092	0.288
Bremen	0.010	0.100	0.010	0.099
Northrhine-Westphalia	0.215	0.411	0.211	0.408
Hesse	0.080	0.271	0.079	0.270
Rhineland-Palatinate	0.043	0.203	0.044	0.204
Baden-Württemberg	0.147	0.354	0.153	0.360

Continued on next page

Table A.2 – continued from previous page

Variable	2014 mean	2014 sd	2018 mean	2018 sd
Bavaria	0.170	0.376	0.171	0.376
Saarland	0.012	0.111	0.012	0.107
Berlin	0.039	0.193	0.041	0.198
Brandenburg	0.025	0.155	0.024	0.154
Mecklenburg Western Pomerania	0.017	0.129	0.016	0.127
Saxony	0.046	0.209	0.045	0.206
Saxony-Anhalt	0.024	0.152	0.022	0.148
Thuringia	0.024	0.153	0.023	0.150
District Type				
Large urban districts	0.354	0.478	0.354	0.478
Urban districts	0.365	0.481	0.366	0.482
Rural districts	0.154	0.361	0.152	0.359
Sparsely populated / rural districts	0.127	0.332	0.128	0.334
Industry (WZ08)				
Agriculture, forestry and fishing	0.009	0.096	0.009	0.092
Mining and quarrying	0.002	0.046	0.002	0.039
Manufacturing	0.221	0.415	0.215	0.411
Electricity, gas, steam and air conditioning supply	0.007	0.085	0.007	0.083
Water supply; sewerage, waste management and remediation activities	0.009	0.092	0.008	0.090
Construction	0.056	0.230	0.054	0.227
Wholesale and retail trade; repair of motor vehicles and motorcycles	0.156	0.363	0.153	0.360
Transportation and storage	0.058	0.234	0.059	0.236
Accommodation and food service activities	0.046	0.210	0.049	0.217
Information and communication	0.032	0.177	0.035	0.184
Financial and insurance activities	0.033	0.177	0.028	0.166
Real estate activities	0.010	0.099	0.010	0.099
Professional, scientific and technical activities	0.066	0.247	0.068	0.252
Administrative and support service activities	0.083	0.275	0.086	0.281
Education	0.020	0.140	0.020	0.139
Human health and social work activities	0.147	0.354	0.152	0.359
Arts, entertainment and recreation	0.013	0.112	0.013	0.114
Other service activities	0.032	0.177	0.031	0.174
Occupation (KldB10, 2-Digit Code)				
Agriculture, animal husbandry and forestry occupations	0.007	0.084	0.007	0.081
Horticultural and floricultural occupations	0.008	0.086	0.008	0.087
Raw material extraction and processing, glass and ceramics production and processing	0.004	0.064	0.004	0.062
Plastics manufacturing and processing, woodworking and wood processing	0.017	0.128	0.016	0.127
Paper and printing occupations, technical media design	0.009	0.096	0.009	0.092
Metal production and processing, metal construction occupations	0.043	0.203	0.041	0.197
Mechanical and automotive engineering occupations	0.053	0.223	0.053	0.224
Mechatronics, energy and electrical occupations	0.031	0.172	0.029	0.167
Technical research, development, design and production control occupations	0.032	0.176	0.034	0.180

Continued on next page

Table A.2 – continued from previous page

Variable	2014 mean	2014 sd	2018 mean	2018 sd
Textile and leather occupations	0.004	0.066	0.004	0.063
Food manufacturing and processing	0.029	0.168	0.029	0.168
Construction planning, architecture and surveying occupations	0.006	0.077	0.007	0.083
Building construction and civil engineering occupations	0.016	0.126	0.017	0.130
(Interior) finishing occupations	0.012	0.108	0.011	0.104
Building and utility engineering occupations	0.021	0.142	0.020	0.141
Mathematics, biology, chemistry and physics occupations	0.012	0.109	0.011	0.105
Geology, geography and environmental protection occupations	0.001	0.037	0.001	0.038
Computer science, information and communication technology occupations	0.021	0.145	0.024	0.153
Transport and logistics occupations (except vehicle driving)	0.066	0.248	0.067	0.250
Vehicle and transport equipment operators	0.038	0.191	0.037	0.188
Protection, security and surveillance occupations	0.010	0.101	0.011	0.102
Cleaning occupations	0.049	0.215	0.045	0.208
Purchasing, distribution and trade occupations	0.030	0.169	0.030	0.170
Sales occupations	0.074	0.261	0.075	0.263
Tourism, hotel and restaurant occupations	0.034	0.181	0.037	0.188
Occupations in business management and organization	0.131	0.337	0.128	0.334
Occupations in financial services, accounting and tax consulting	0.046	0.209	0.042	0.199
Professions in law and administration	0.014	0.118	0.014	0.117
Medical health professions	0.079	0.269	0.080	0.271
Non-medical health, personal care and wellness occupations, medical technology	0.028	0.164	0.030	0.170
Education, social and domestic professions, theology	0.043	0.202	0.046	0.210
Teaching and training occupations	0.011	0.103	0.010	0.100
Linguistic, literary, humanistic, social and economic professions	0.002	0.044	0.003	0.051
Advertising, marketing, commercial and editorial media occupations	0.016	0.126	0.018	0.132
Product design and arts and crafts occupations, fine arts, musical instrument making	0.002	0.044	0.002	0.044
Performing and entertainment occupations	0.004	0.063	0.004	0.062
Union coverage				
No coverage	0.615	0.487	0.636	0.481
Sectoral agreement	0.310	0.462	0.293	0.455
Plant / firm agreement	0.043	0.204	0.042	0.200
Company agreement	0.032	0.177	0.030	0.170
Participation of the public sector				
Yes	0.072	0.259	0.063	0.243
Firmsize				
< 10 empl.	0.152	0.359	0.141	0.348
10 to 49 empl.	0.244	0.430	0.246	0.431
50 to 99 empl	0.106	0.307	0.104	0.305
100 to 249 empl.	0.133	0.340	0.139	0.346
250 to 499 empl.	0.094	0.291	0.095	0.293
500 to 1,000 empl.	0.075	0.263	0.074	0.262
> 1,000 empl.	0.197	0.398	0.200	0.400

Continued on next page

Table A.2 – continued from previous page

	2014		2018	
Variable	mean	sd	mean	sd
Sectoral minimum wage				
Yes	0.247	0.431	0.244	0.430
No	0.599	0.490	0.612	0.487
Unknown	0.154	0.361	0.144	0.351
Number of observations	708,081		693,827	

Notes: Survey weights have been used for all calculations.
Source: GSES 2014 and 2018 and own calculations.

A.7 DGUV-IAB data

A.7.1 Adjustment of hours worked in the DGUV-IAB data

Typically, employers directly report working hours to the DGUV (Deutsche Gesetzliche Unfallversicherung/German Social Accident Insurance) for administrative purposes. However, in the years 2011 to 2014, the reporting of working hours was conducted through the social security notification system, allowing us to link this information to other administrative IAB data. Our goal is to construct from this information weekly hours worked and hourly wages that are as close as possible to those defined in the GSES. Wage earnings in administrative data are always reported as the average daily wage, i.e., total spell earnings divided by the spell length in days. As a consequence, weekly wages are calculated as the daily wage times seven week days. Unfortunately, when reporting working hours for an employment spell, employers could either report *actual hours worked* (excluding annual and sick leave) or *contractual hours worked* (referring to calendar periods, i.e., including periods of leave). In addition, employers could report 'educated guesses'. It is not evident in the data which reporting mode was selected (vom Berge et al., 2014).

Clearly, it matters for the correct calculation of weekly working hours whether employers based their reports on the number of days including or excluding times of leave. We therefore followed heuristics similar to Dustmann et al. (2022) to calculate an adequate number of hours worked per week. In a first step, we classified notifications into whether they were more likely of the 'actual hours' or the 'contractual hours' type. The notifications of the 'contractual hours' type did not need to be adjusted as they correspond to calendar days and weeks thus directly referring to the spell length. By contrast, total spell hours in the 'actual hours' reporting type would result in a too low number of weekly hours if one were to divide them by spell weeks as the count of total spell hours excludes times of leave. We therefore use the same adjustment factor of 1.19 as in Dustmann et al. (2022) to upward-correct weekly hours in these cases. The adjustment factor 1.19 is computed as 250/210, i.e., the ratio of potential working days and working days net of average annual and sick leave. For cases not classified into 'actual' or 'contractual hours' by the heuristics below, we do not make any adjustment as this gives the best fit of the resulting distribution of weekly hours and hourly wages to those observed in the GSES (see descriptive statistics below). Finally, following Dustmann et al. (2022), we add to the value of weekly hours worked obtained by the procedure above typical values for overtime in order to align the weekly hours count to that in the GSES (1.24 hours for full-time employees, 0.56 hours for part-time employees and 0.19 hours for marginal part-time workers).

Heuristic 1 – '*Vollarbeitsrichtwerte*'

If an employment spell was an annual spell and the reported total hours worked was identical to the so-called 'Vollarbeitsrichtwert' (standard annual full-time equivalent, see vom Berge et al., 2014), we classified this notification as 'actual hours' since 'Vollarbeitsrichtwerte' are computed net of annual and sick leave. 'Vollarbeitsrichtwerte' are published on a regular basis and employers may refer to them when reporting working hours. As a consequence, we applied the adjustment factor to these kind of observations.

Heuristic 2 – '*Conspicuous values when dividing by number of weeks*'

If dividing total spell hours by spell weeks resulted in the conspicuously exact values 35, 35.5, 36, ..., 40.5, 41, we concluded that employers constructed their reported total hours by multiplying the number of weeks by typical values for weekly working hours. In this case, it is clear that employers refer to 'contractual hours' as the weeks include times of leave.

Similarly, if we obtained the above exact values 35, 35.5, 36, ..., 40.5, 41 when dividing total spell hours by 45 or 46 weeks (which would be the typical number of annual weeks worked excluding annual leave times), we concluded that employers refer to 'actual hours' and thus applied the adjustment factor of 1.19.

Heuristic 3 – '*Heaping*'

Apart from the cases above, we observed conspicuous heaping at a number of values for reported total hours (relative frequency at least 5%). If dividing these values by 52 in the case of annual spells (most spells are annual spells) led to weekly working hours below 35 hours for full-time employees, we concluded that employers were of the 'actual hours' reporting type and thus applied the adjustment factor (as assuming they were of the ‘contractual hours’ type would lead to a contradiction with the information that a given worker works full-time).

Heuristic 4 – ‘*Firm-level behavior*’

Following Dustmann et al. (2022), we divide total spell hours by spell weeks. If this results in weekly working hours below 35 hours for at least 80% of the employees of an employer, we classify this employer as being of the ‘actual hours’ reporting type and apply the adjustment factor of 1.19 to all its employees.

A.7.2 Covariates not available in DGUV-IAB data

The following covariates are available in the GSES but not in the DGUV-IAB data: whether the firm was associated to the public sector, whether the firm was covered by a sectoral minimum wage, union coverage, worker tenure. Our main results are insensitive to including/excluding these variables in the GSES analysis.

A.7.3 Descriptive statistics GSES vs. DGUV-IAB

Table A.3: Hourly wages

	Hourly wages	
	GSES	DGUV-IAB
Percentile		
5	7.461	6.781
10	8.464	8.211
15	9.079	9.249
20	9.709	10.230
25	10.389	11.245
30	11.266	12.276
35	12.119	13.311
40	12.974	14.395
45	13.835	15.486
50	14.698	16.643
55	15.626	17.869
60	16.654	19.174
65	17.811	20.593
70	19.145	22.226
75	20.698	24.183
80	22.720	26.612
85	25.411	29.712
90	29.592	33.232

Sources: DGUV-IAB (2011-2014), GSES (2014)

Table A.4: Weekly hours worked

	Relative frequencies	
	GSES	DGUV-IAB
Threshold		
7	0.055	0.038
12	0.077	0.070
20	0.076	0.088
25	0.054	0.057
30	0.046	0.042
35	0.099	0.212
42	0.531	0.417
51	0.055	0.052
Max	0.007	0.024

Sources: DGUV-IAB (2011-2014), GSES (2014)

A.8 Additional results

A.8.1 Hourly wages

Table A.5: Minimum wage effects on inequality in hourly wages, 2014 vs. 2018 – Full controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	16.247 (0.241)	0.260 (0.002)	3.504 (0.073)	2.015 (0.031)	1.739 (0.012)
2018	17.740 (0.208)	0.240 (0.002)	3.339 (0.080)	1.992 (0.030)	1.676 (0.018)
$\hat{\Delta}^{18-14}$	1.493*** (0.329)	−0.020*** (0.003)	−0.165 (0.113)	−0.023 (0.045)	−0.063*** (0.022)
Bite 1 (Regions)					
No trend adjustment	0.371*** (0.064)	−0.035*** (0.002)	−0.829*** (0.066)	−0.118*** (0.032)	−0.299*** (0.023)
1-year trend adjustment	0.309*** (0.066)	−0.029*** (0.002)	−0.670*** (0.063)	−0.116*** (0.032)	−0.225*** (0.023)
2-year trend adjustment	0.233*** (0.074)	−0.022*** (0.002)	−0.529*** (0.060)	−0.114*** (0.033)	−0.160*** (0.022)
Bite 2 (Augmented occupations)					
No trend adjustment	0.032 (0.071)	−0.027*** (0.006)	−0.750*** (0.109)	−0.050 (0.038)	−0.326*** (0.041)
1-year trend adjustment	0.053 (0.071)	−0.022*** (0.005)	−0.635*** (0.102)	−0.057 (0.038)	−0.263*** (0.037)
2-year trend adjustment	0.056 (0.112)	−0.017*** (0.005)	−0.534*** (0.106)	−0.062 (0.045)	−0.209*** (0.036)
Bite 3 (Augmented industries)					
No trend adjustment	0.075* (0.043)	−0.026*** (0.004)	−0.762*** (0.118)	−0.057 (0.042)	−0.325*** (0.052)
1-year trend adjustment	0.085* (0.047)	−0.021*** (0.003)	−0.638*** (0.119)	−0.064 (0.045)	−0.258*** (0.051)
2-year trend adjustment	0.088 (0.067)	−0.017*** (0.003)	−0.502*** (0.118)	−0.070 (0.053)	−0.186*** (0.051)

Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Table A.6: Minimum wage effects on inequality in hourly wages, 2014 vs. 2018 – No firm controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	16.247 (0.241)	0.260 (0.002)	3.504 (0.073)	2.015 (0.031)	1.739 (0.012)
2018	17.740 (0.208)	0.240 (0.002)	3.339 (0.080)	1.992 (0.030)	1.676 (0.018)
$\hat{\Delta}^{18-14}$	1.493*** (0.329)	−0.020*** (0.003)	−0.165 (0.113)	−0.023 (0.045)	−0.063*** (0.022)
Bite 1 (Regions)					
No trend adjustment	0.389*** (0.067)	−0.036*** (0.002)	−0.826*** (0.063)	−0.119*** (0.030)	−0.296*** (0.023)
1-year trend adjustment	0.330*** (0.071)	−0.029*** (0.002)	−0.668*** (0.060)	−0.118*** (0.030)	−0.223*** (0.022)
2-year trend adjustment	0.256*** (0.081)	−0.023*** (0.002)	−0.529*** (0.057)	−0.117*** (0.031)	−0.158*** (0.022)
Bite 2 (Augmented occupations)					
No trend adjustment	0.049 (0.075)	−0.027*** (0.006)	−0.752*** (0.109)	−0.054 (0.038)	−0.324*** (0.041)
1-year trend adjustment	0.064 (0.075)	−0.022*** (0.005)	−0.636*** (0.102)	−0.060 (0.037)	−0.261*** (0.037)
2-year trend adjustment	0.067 (0.113)	−0.017*** (0.005)	−0.535*** (0.106)	−0.064 (0.044)	−0.208*** (0.036)
Bite 3 (Augmented industries)					
No trend adjustment	0.077 (0.047)	−0.026*** (0.004)	−0.763*** (0.117)	−0.058 (0.041)	−0.325*** (0.052)
1-year trend adjustment	0.087* (0.050)	−0.021*** (0.003)	−0.639*** (0.117)	−0.065 (0.044)	−0.257*** (0.051)
2-year trend adjustment	0.089 (0.069)	−0.017*** (0.003)	−0.503*** (0.116)	−0.071 (0.052)	−0.186*** (0.051)

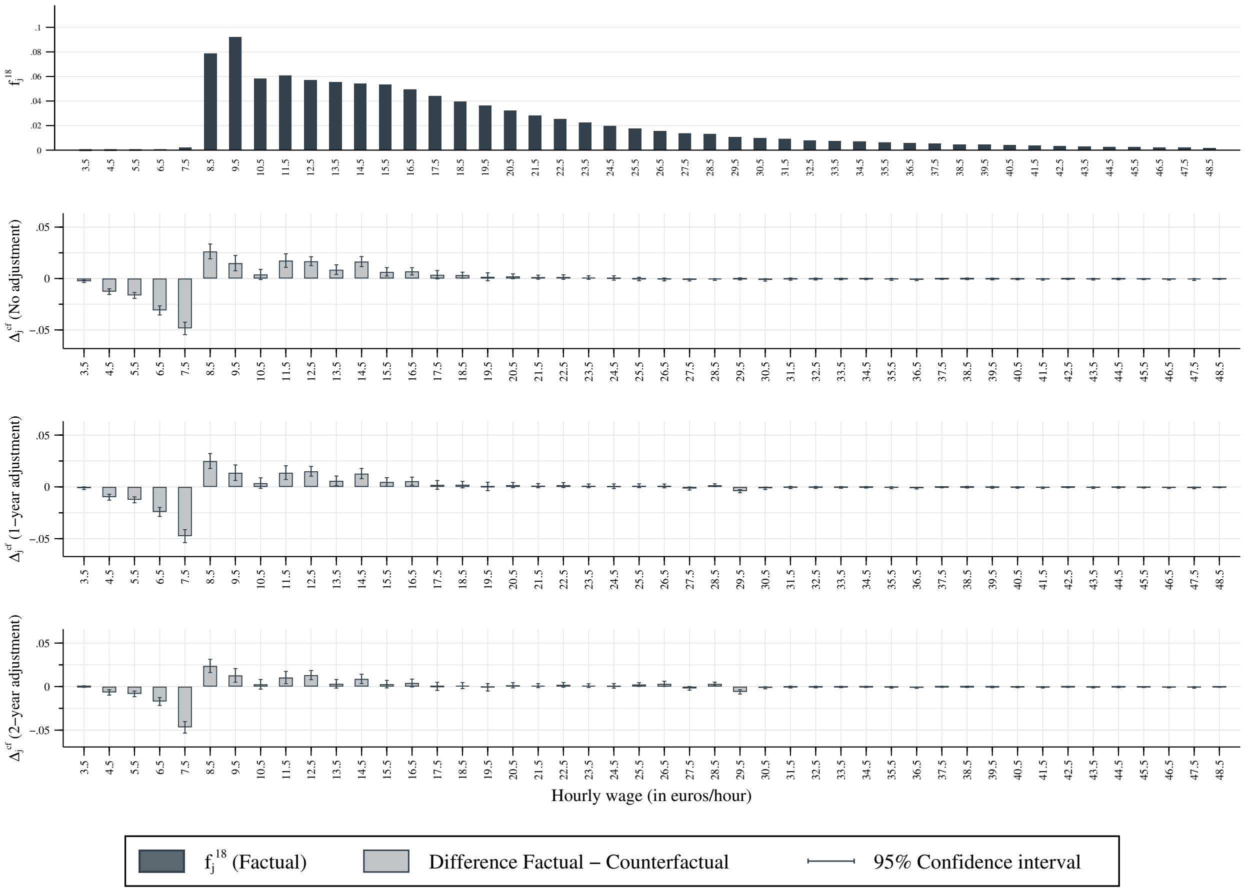
Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Table A.7: Minimum wage effects on inequality in hourly wages, 2014 vs. 2018 – No additional controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	16.247	0.260	3.504	2.015	1.739
	(0.241)	(0.002)	(0.073)	(0.031)	(0.012)
2018	17.740	0.240	3.339	1.992	1.676
	(0.208)	(0.002)	(0.080)	(0.030)	(0.018)
$\hat{\Delta}^{18-14}$	1.493***	−0.020***	−0.165	−0.023	−0.063***
	(0.329)	(0.003)	(0.113)	(0.045)	(0.022)
Bite 1 (Regions)					
No trend adjustment	0.513***	−0.037***	−0.830***	−0.137***	−0.282***
	(0.104)	(0.003)	(0.065)	(0.035)	(0.024)
1-year trend adjustment	0.459***	−0.030***	−0.678***	−0.134***	−0.213***
	(0.106)	(0.002)	(0.063)	(0.035)	(0.023)
2-year trend adjustment	0.391***	−0.024***	−0.547***	−0.131***	−0.154***
	(0.117)	(0.002)	(0.064)	(0.037)	(0.021)
Bite 2 (Augmented occupations)					
No trend adjustment	0.068	−0.028***	−0.778***	−0.064	−0.327***
	(0.084)	(0.006)	(0.114)	(0.042)	(0.039)
1-year trend adjustment	0.092	−0.024***	−0.665***	−0.071*	−0.264***
	(0.082)	(0.005)	(0.106)	(0.041)	(0.034)
2-year trend adjustment	0.099	−0.019***	−0.567***	−0.076	−0.212***
	(0.115)	(0.005)	(0.110)	(0.047)	(0.033)
Bite 3 (Augmented industries)					
No trend adjustment	0.121*	−0.028***	−0.773***	−0.065	−0.323***
	(0.064)	(0.004)	(0.113)	(0.043)	(0.050)
1-year trend adjustment	0.137**	−0.023***	−0.656***	−0.072	−0.259***
	(0.068)	(0.004)	(0.117)	(0.046)	(0.050)
2-year trend adjustment	0.143*	−0.018***	−0.529***	−0.079	−0.192***
	(0.077)	(0.003)	(0.118)	(0.053)	(0.049)

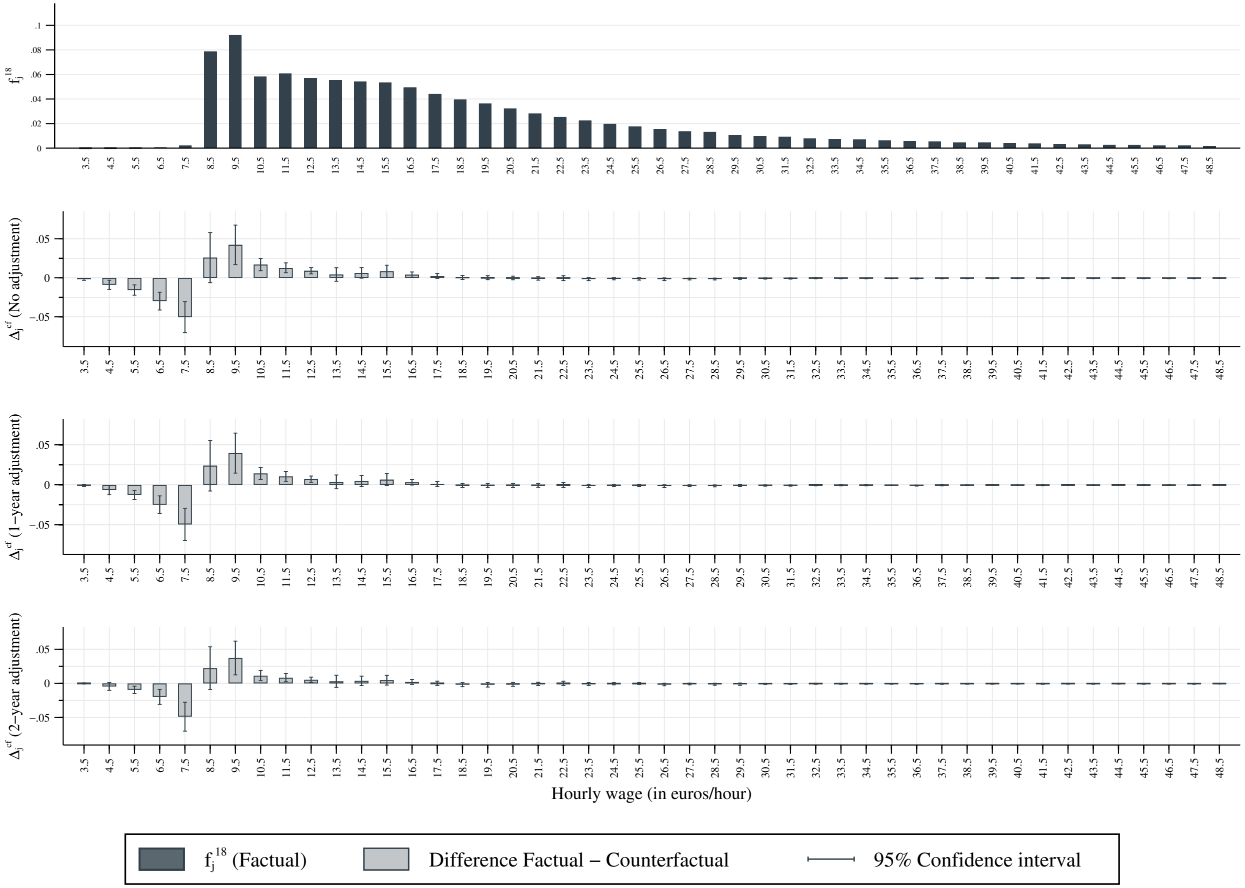
Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.2: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend trend adjustment regimes. Bite 1: Regions. Full controls.



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

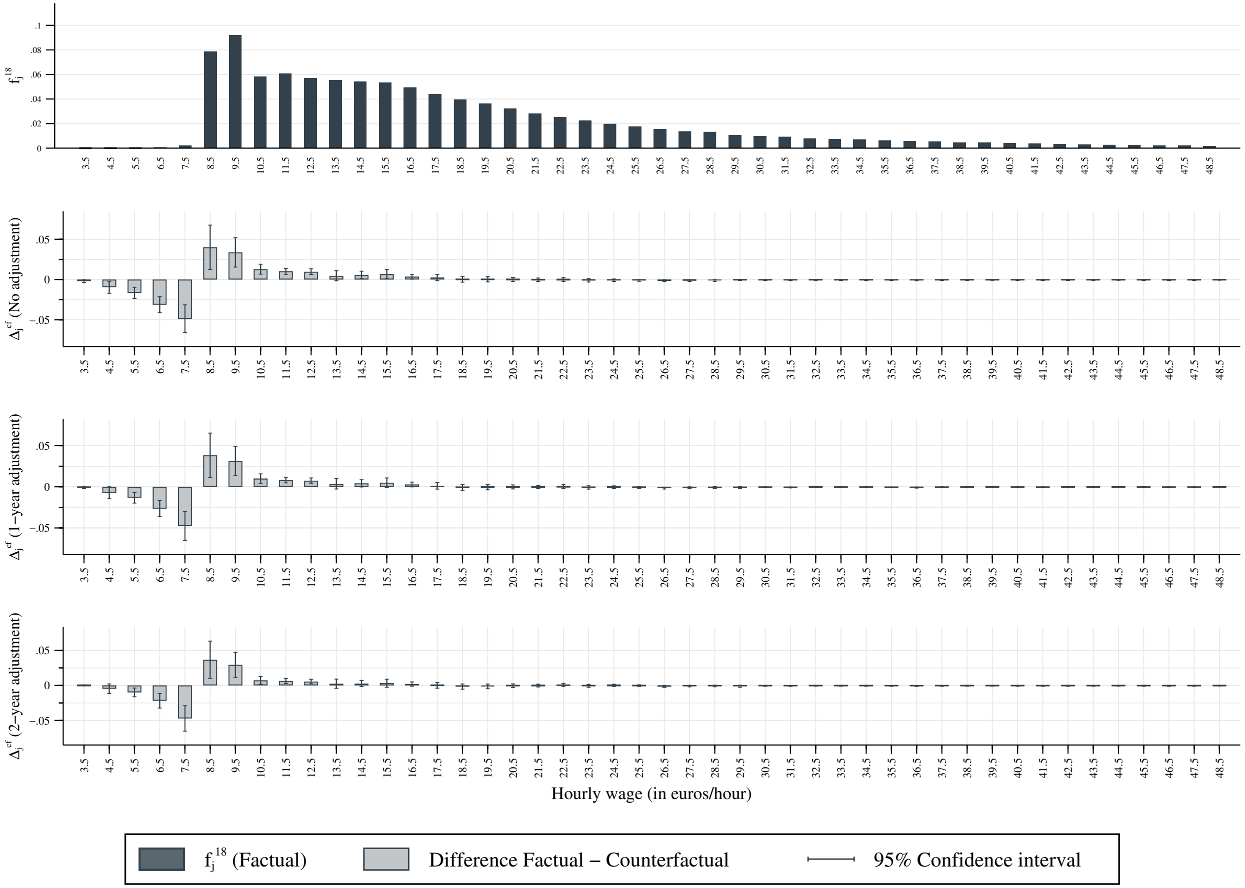
Figure A.3: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. Full controls.



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

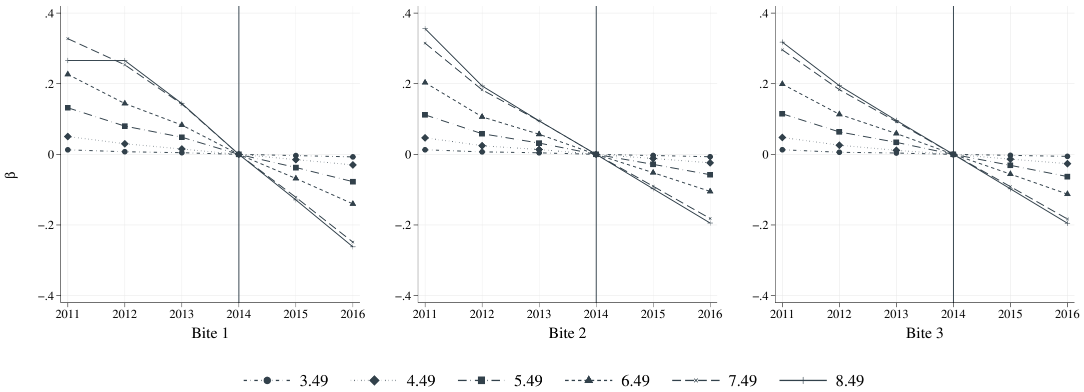
Figure A.4: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. Full controls.



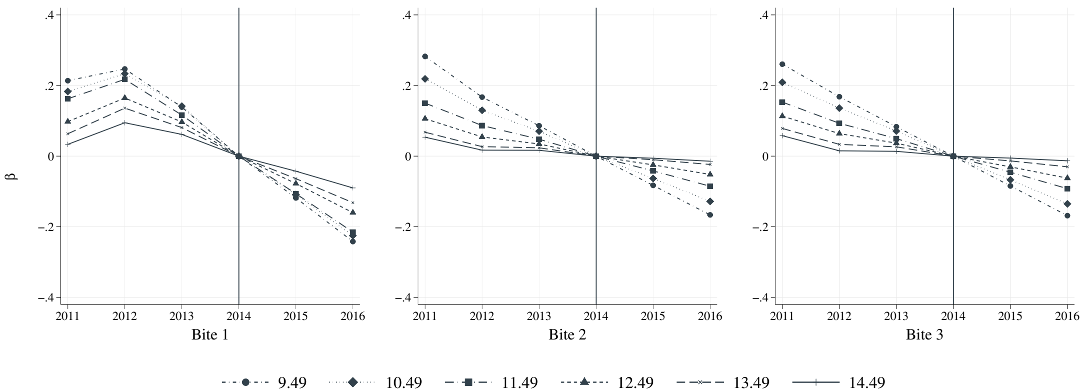
Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.5: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Hourly wages, all bites. Full controls.



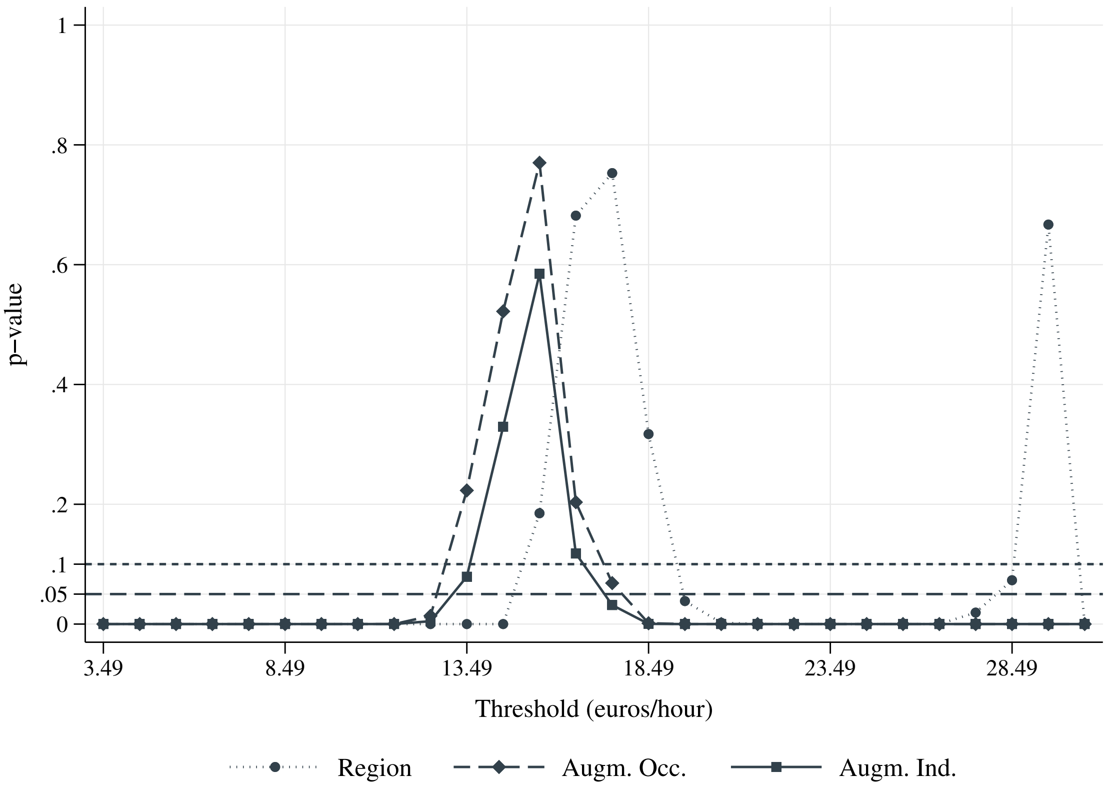
(a) Lower thresholds (3.49 to 8.49)



(b) Upper thresholds (9.49 to 14.49)

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for bins below and above the minimum wage level. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.6: P-values of joint significance – Hourly wages. Full controls.

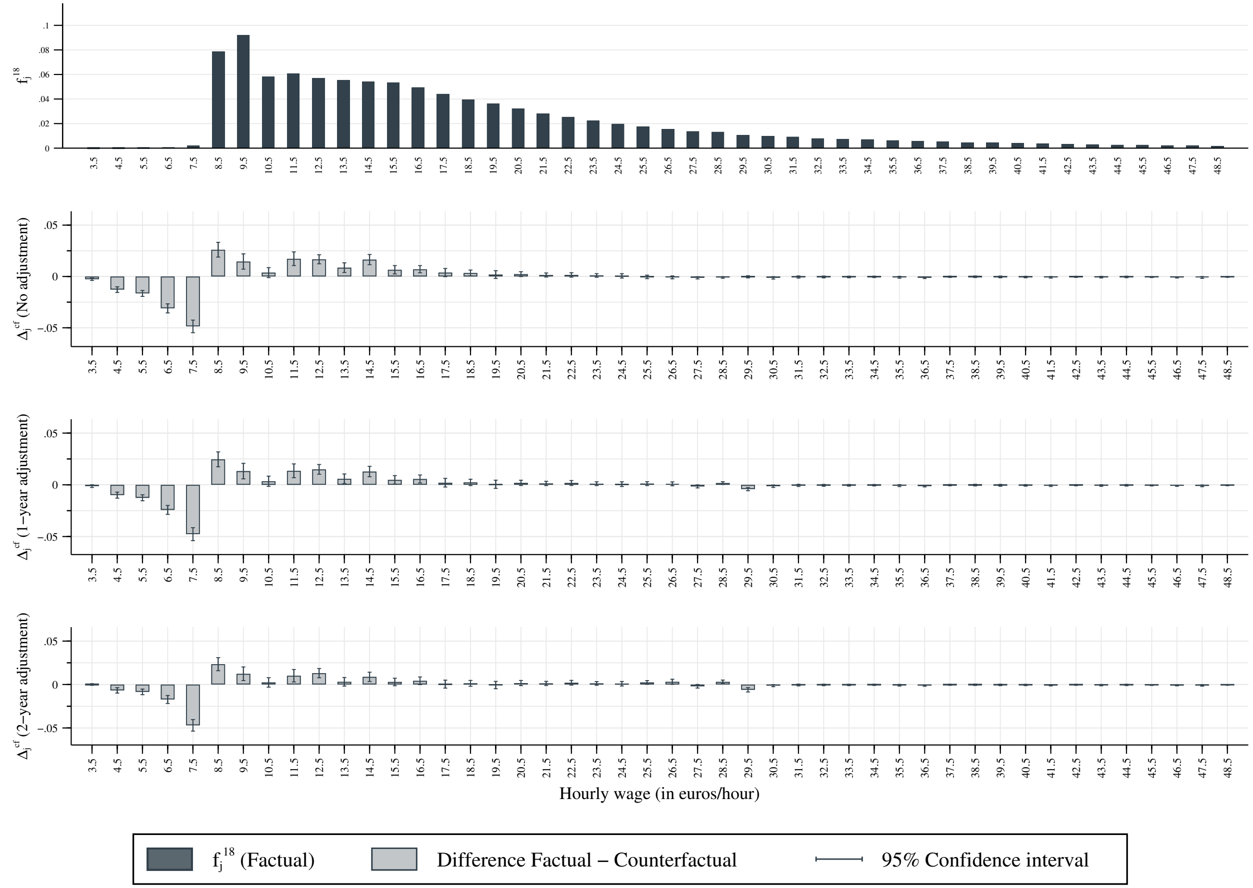


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0 : \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No firm controls

— Intentionally left blank —

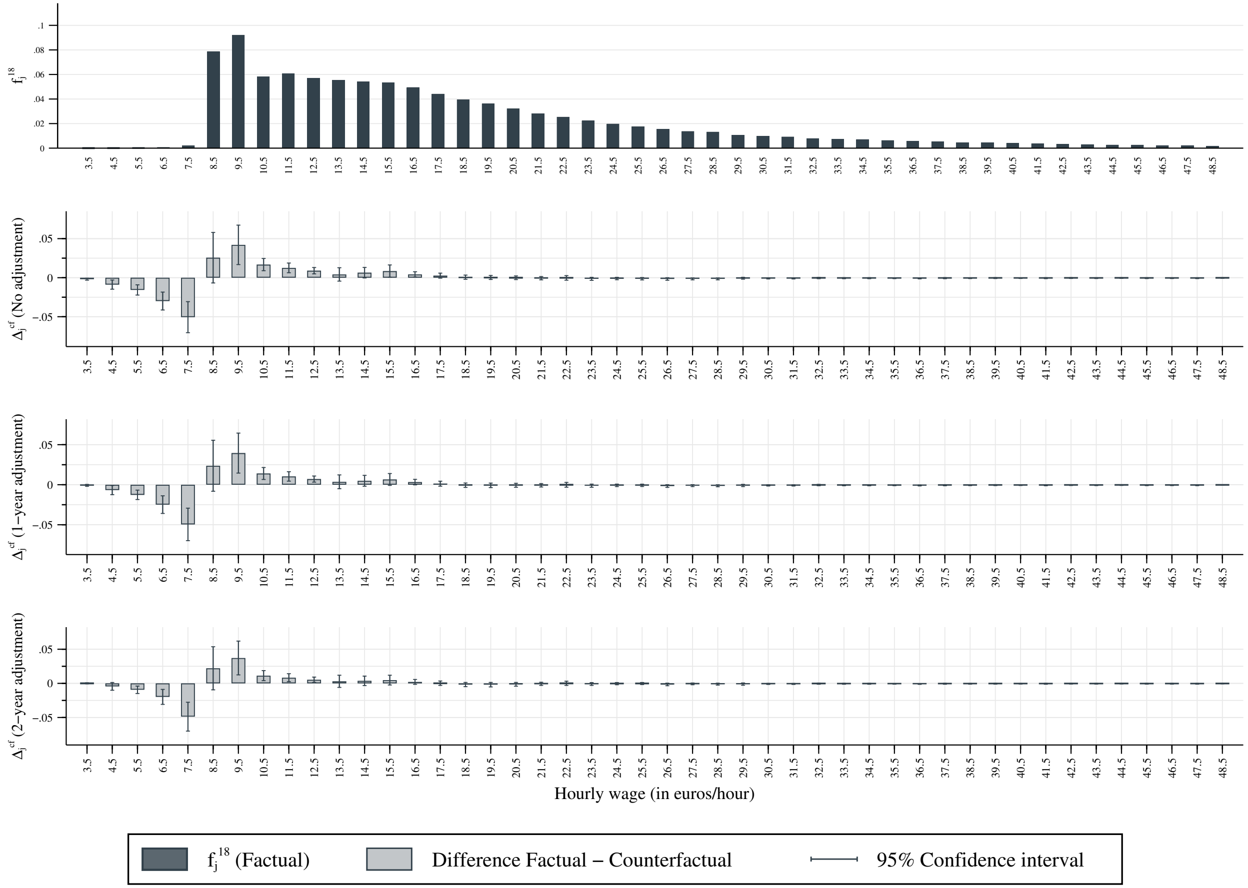
Figure A.7: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 1: Regions. No firm controls.



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

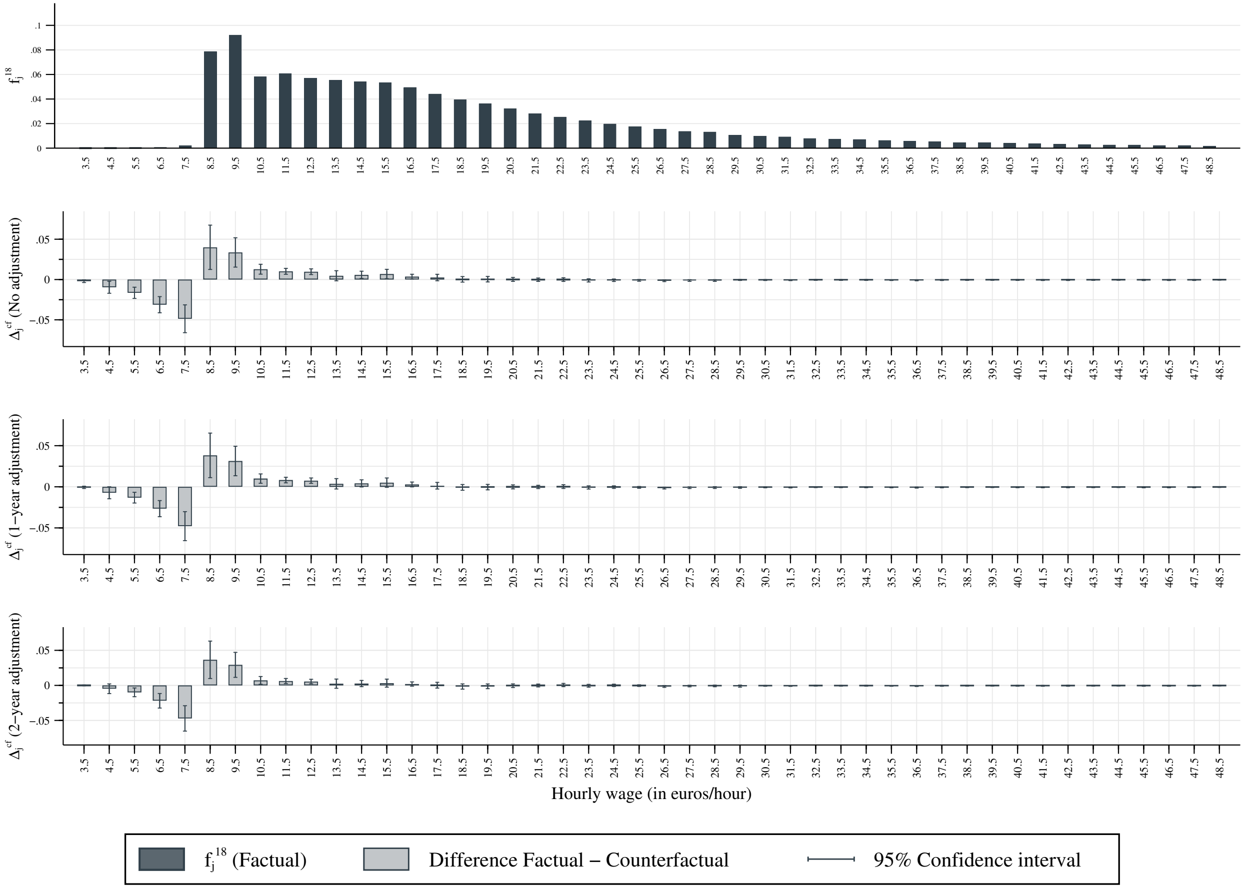
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.8: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. No firm controls.



Notes: The x-axis shows hourly wage bins. For example, the ‘10.50’ bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

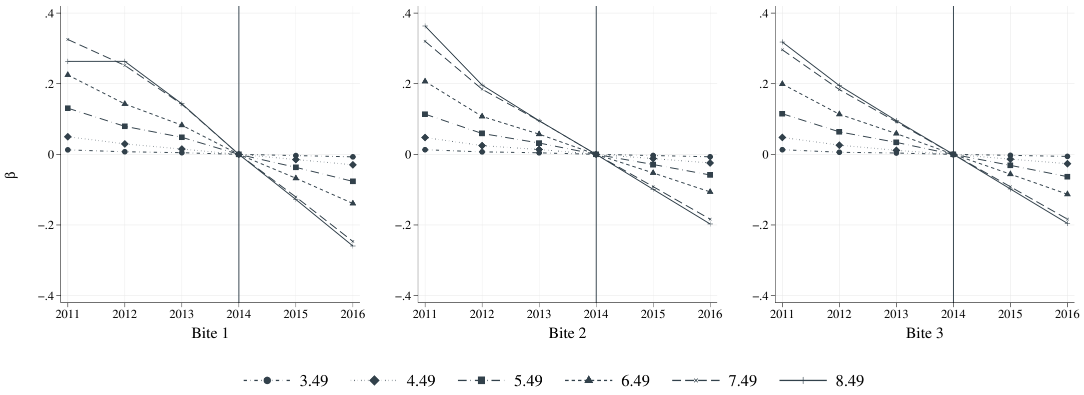
Figure A.9: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. No firm controls.



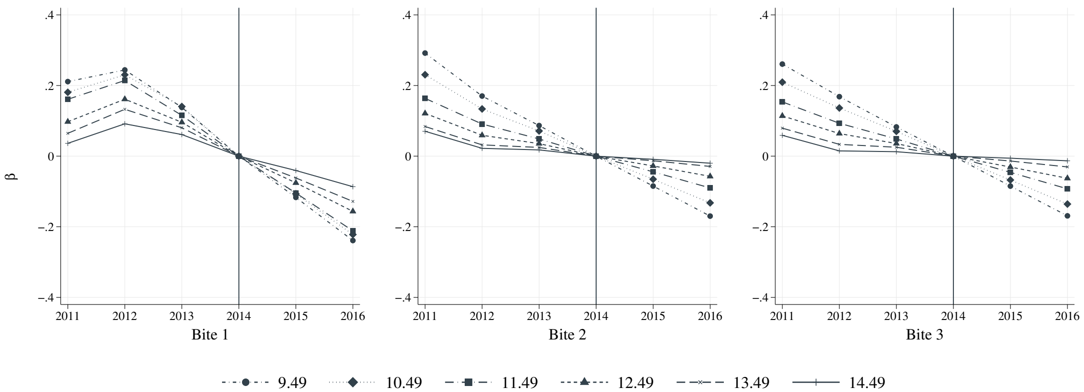
Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.10: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Hourly wages, all bites. No firm controls.



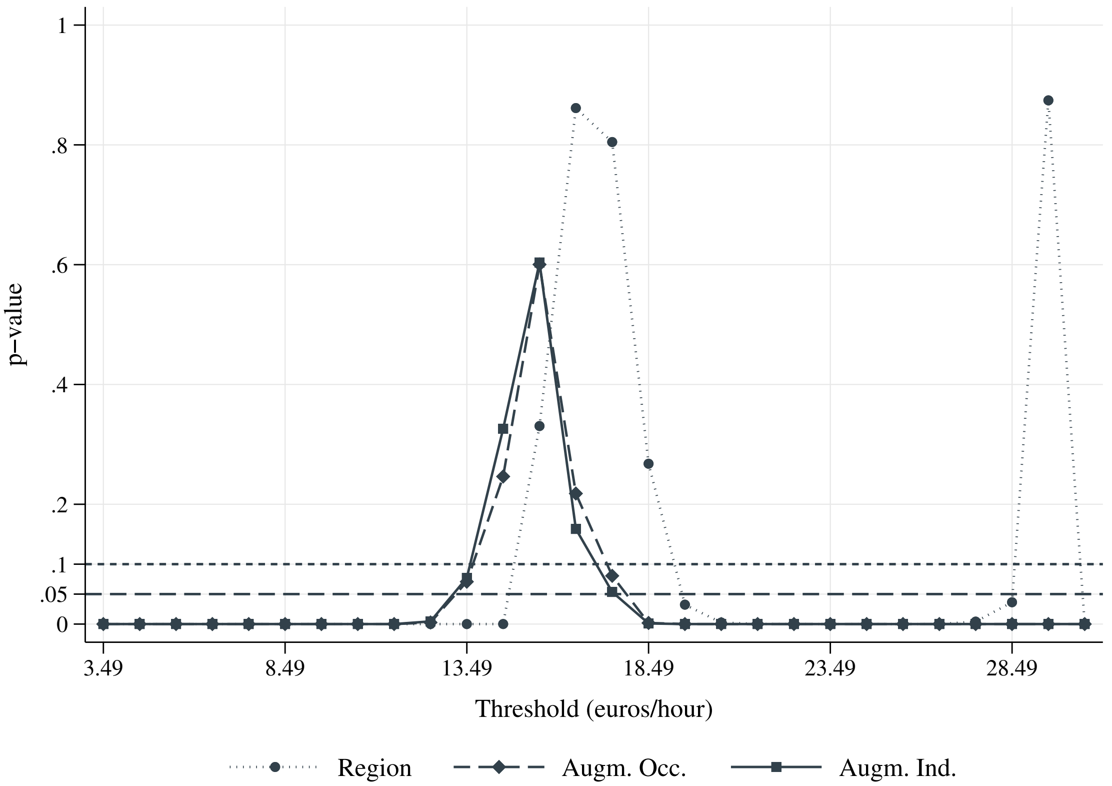
(a) Lower thresholds (3.49 to 8.49)



(b) Upper thresholds (9.49 to 14.49)

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for bins below and above the minimum wage level. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.11: P-values of joint significance – Hourly wages. No firm controls.

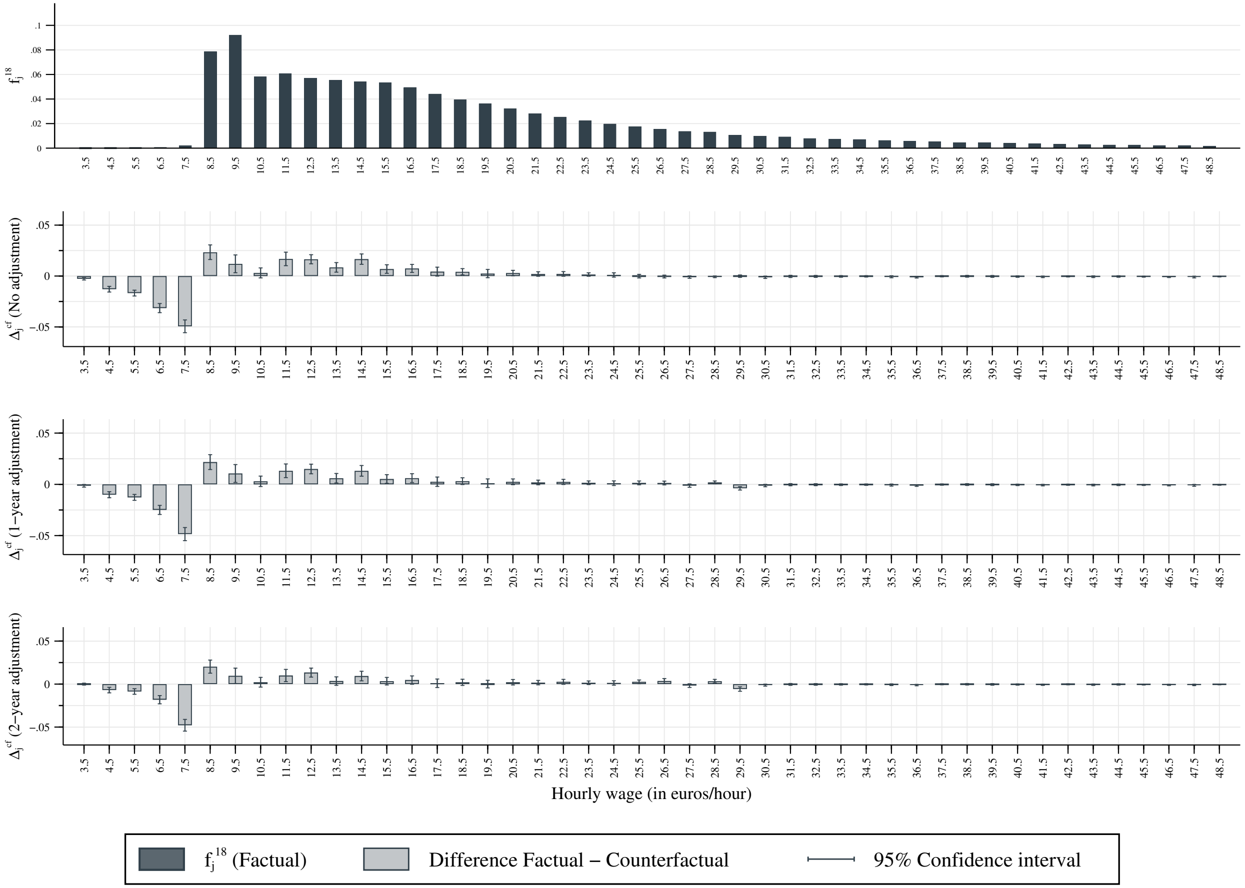


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0 : \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No controls

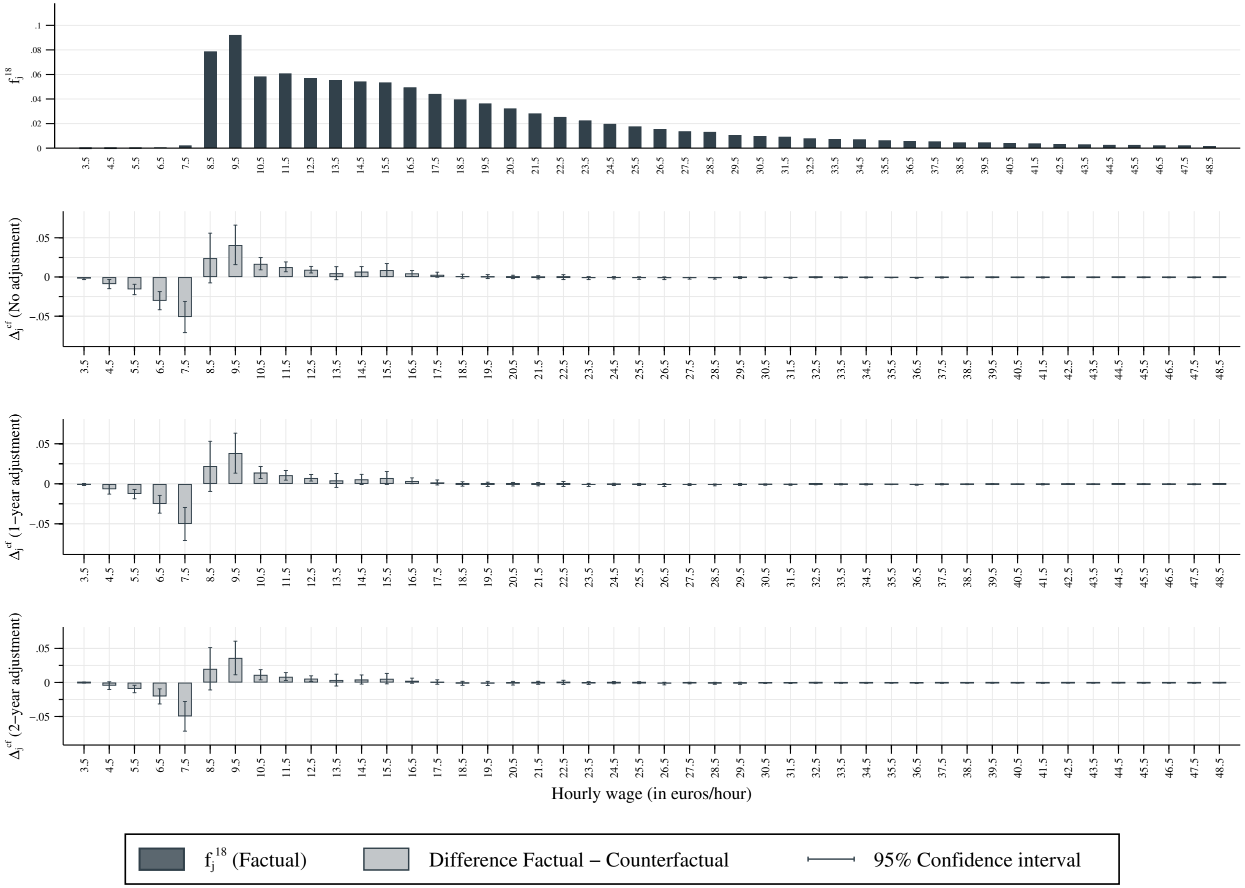
— Intentionally left blank —

Figure A.12: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 1: Regions. No controls.



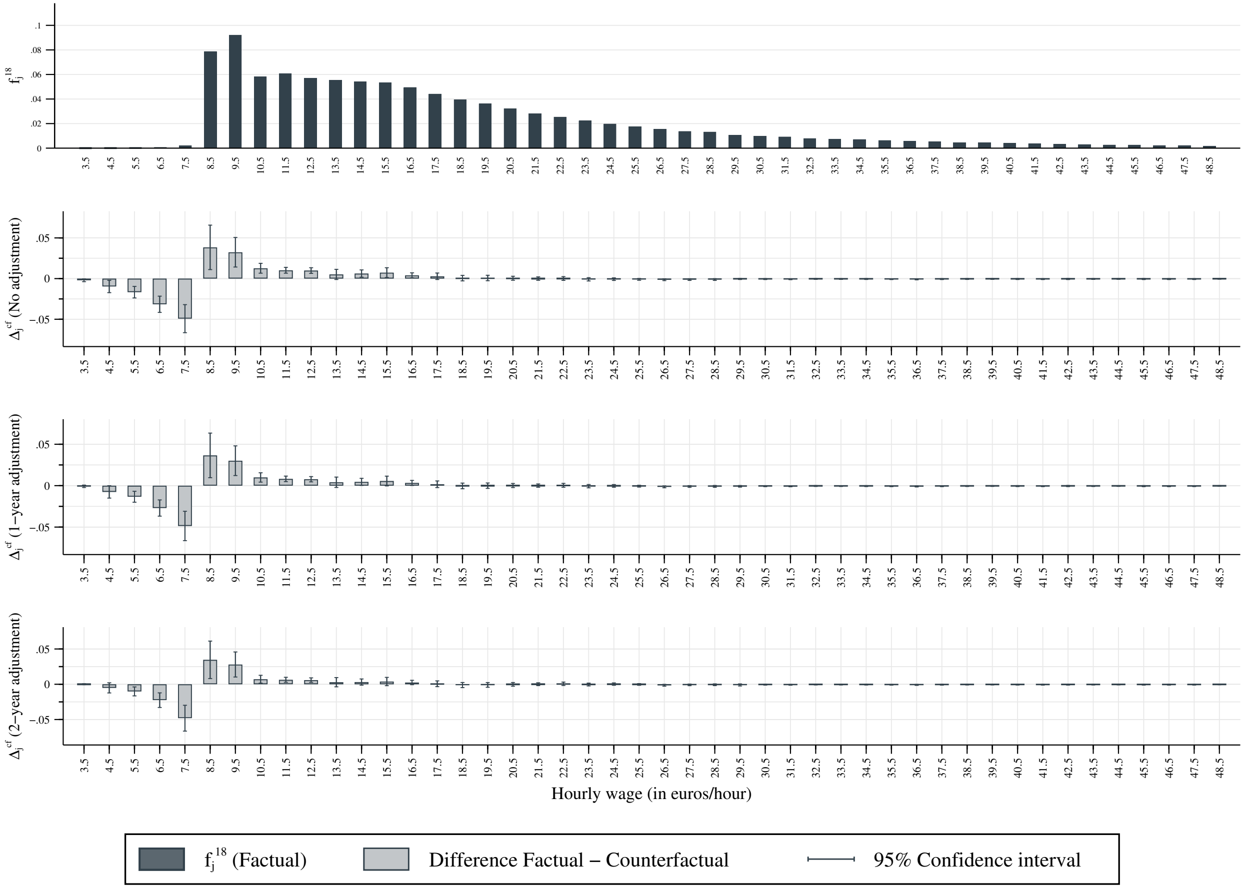
Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.13: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. No controls.



Notes: The x-axis shows hourly wage bins. For example, the '10.50' bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

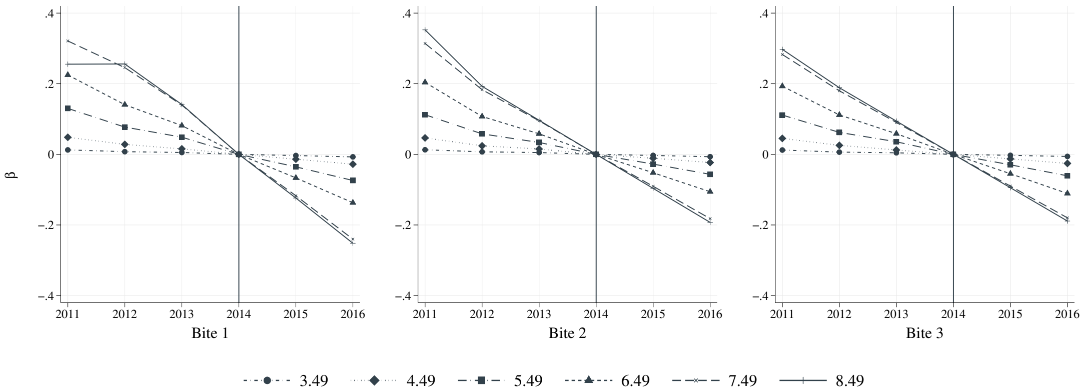
Figure A.14: 2018 Factual hourly wages distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. No controls.



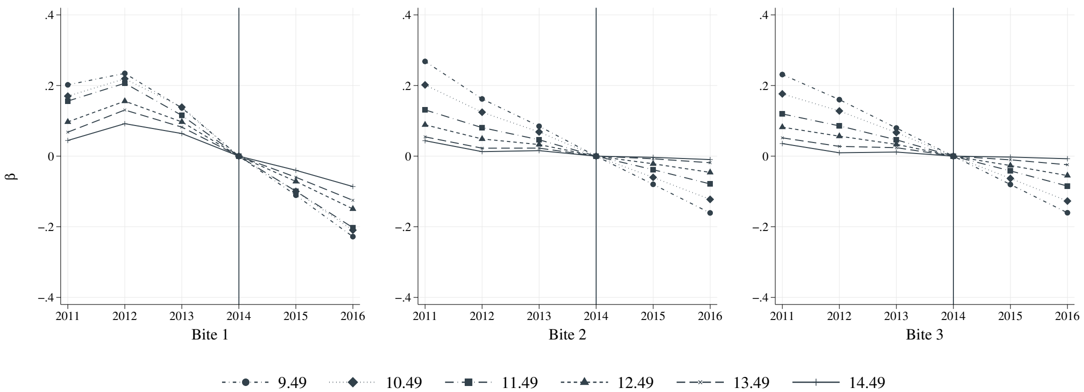
Notes: The x-axis shows hourly wage bins. For example, the ‘10.50’ bin comprises hourly wages in the interval [10.50; 11.49] euros/hour. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.15: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Hourly wages, all bites. No controls.



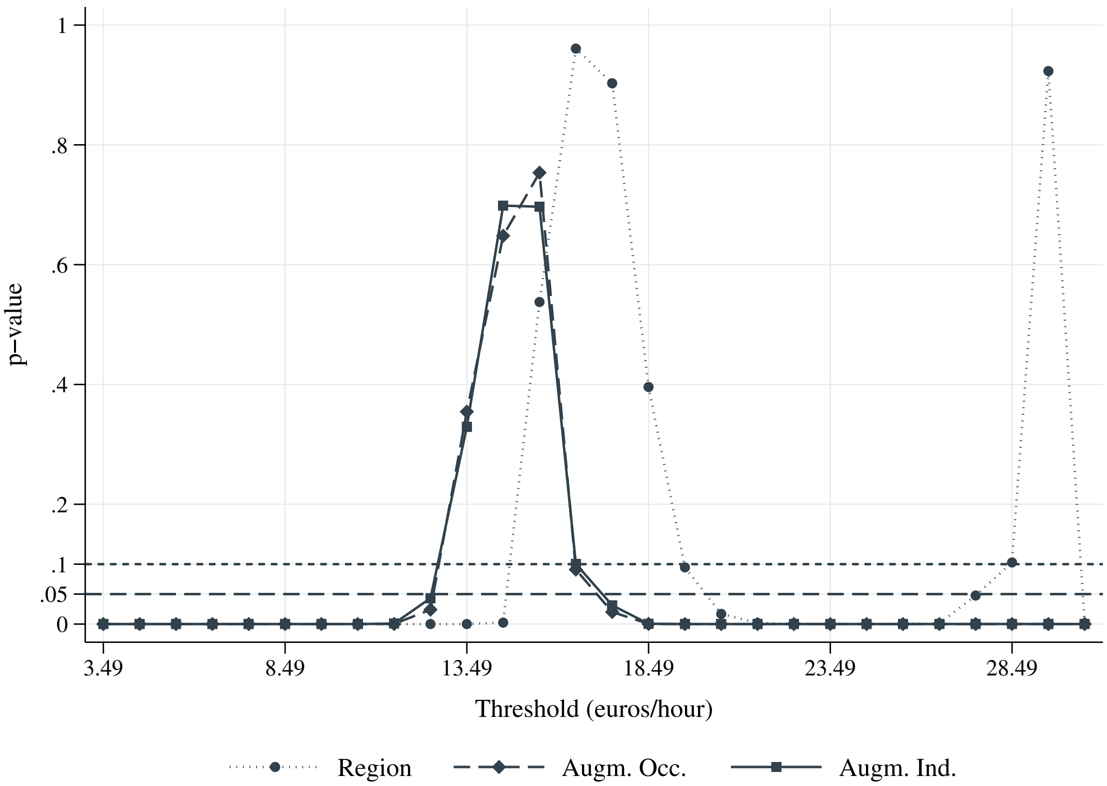
(a) Lower thresholds (3.49 to 8.49)



(b) Upper thresholds (9.49 to 14.49)

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for bins below and above the minimum wage level. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.16: P-values of joint significance – Hourly wages. No controls.



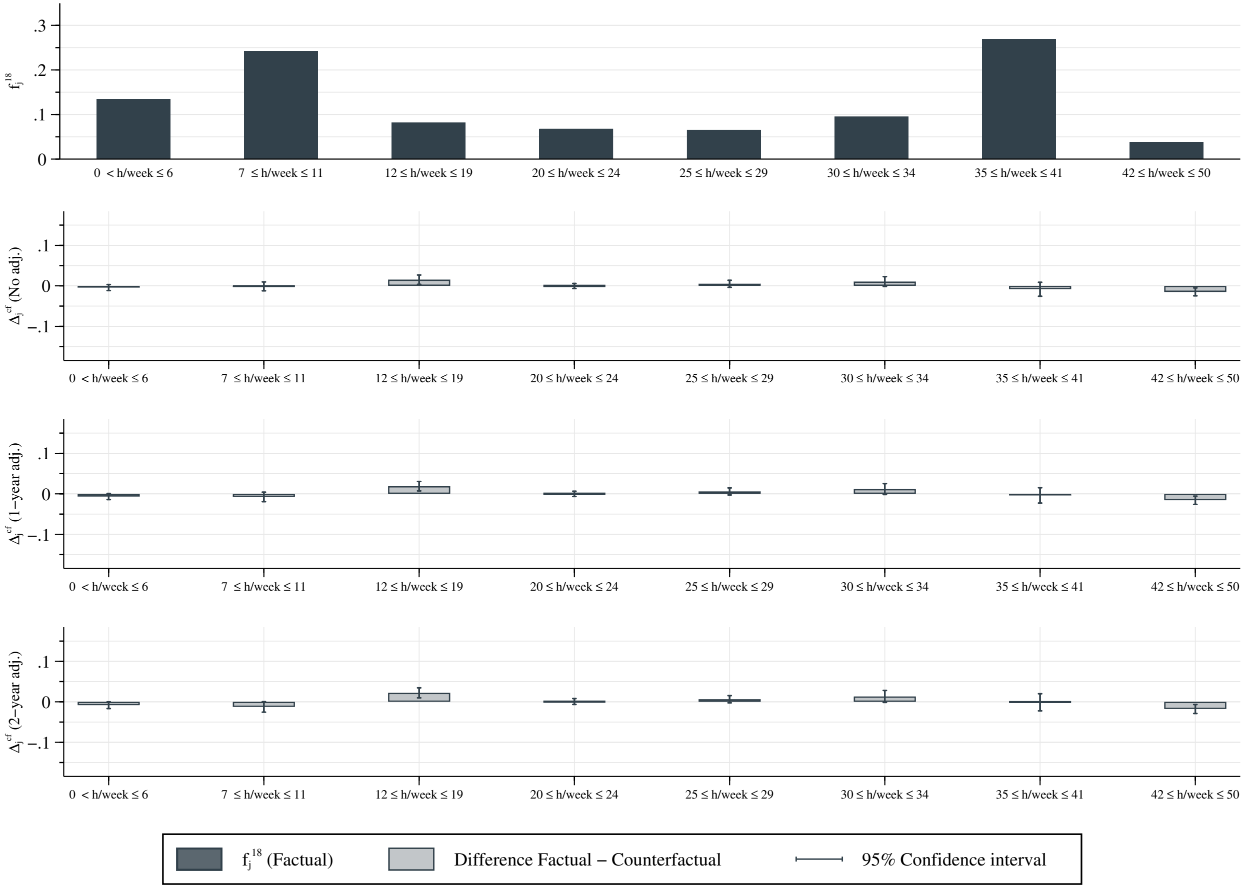
Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

A.8.2 Weekly hours worked

Hourly wages ≤ 12 euros/hour

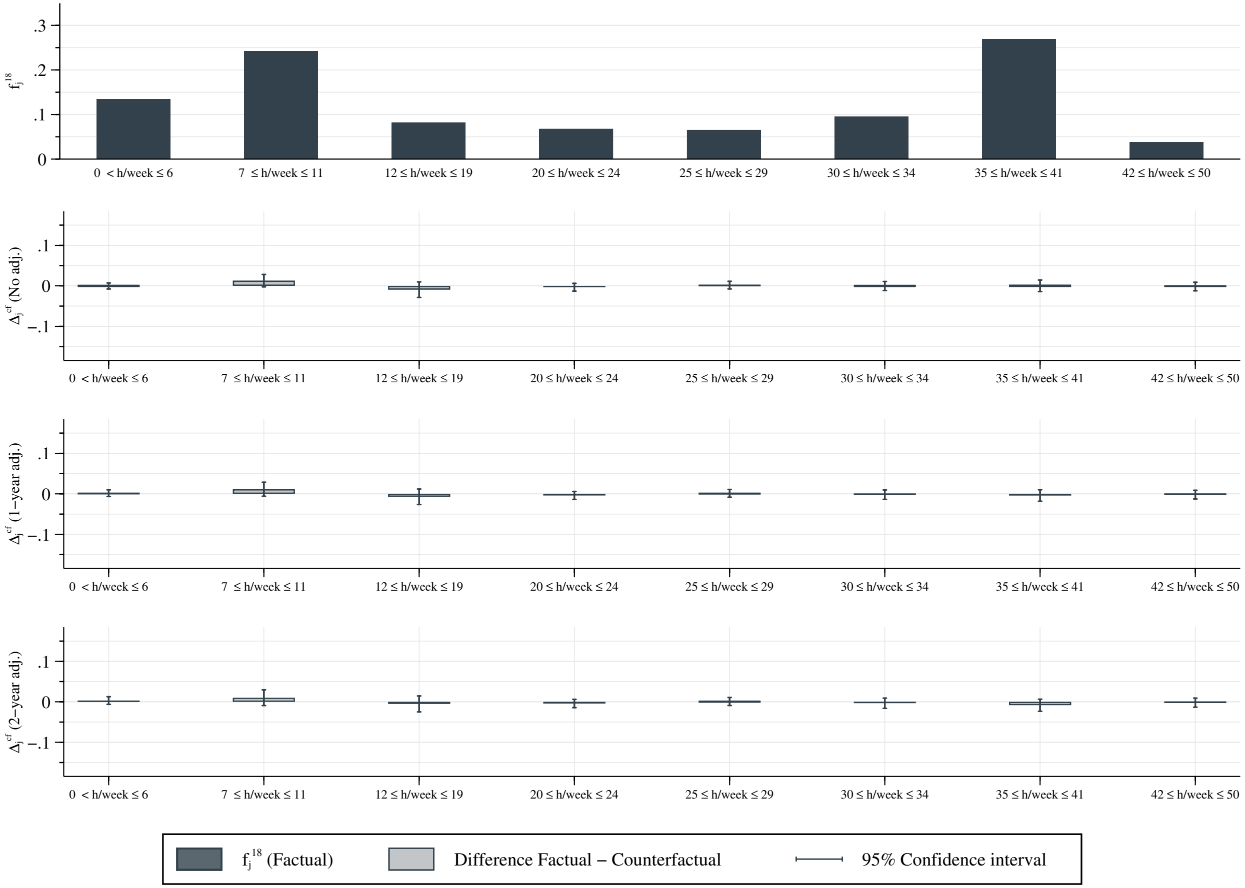
— Intentionally left blank —

Figure A.17: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 1: Regions. All trend adjustment regimes. Full controls.



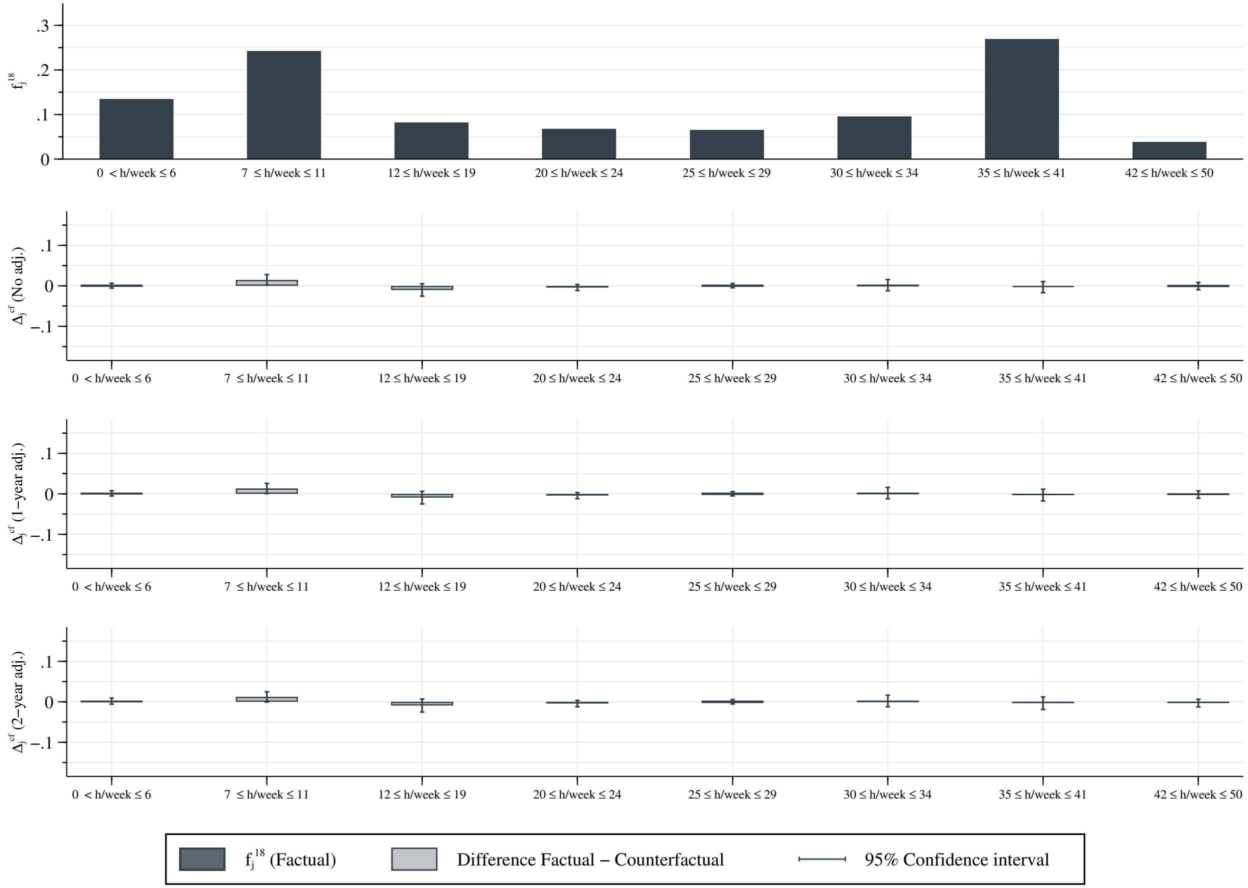
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.18: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. Full controls.



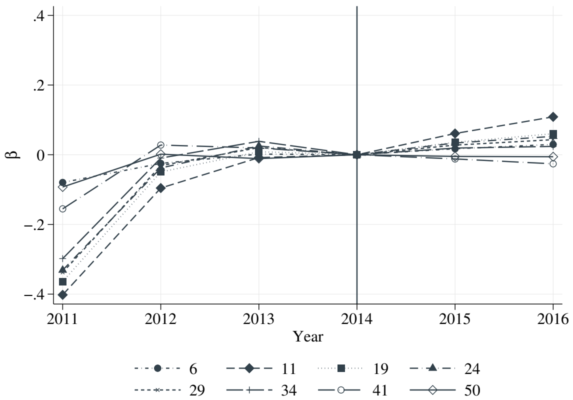
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.19: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages ≤ 12 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. Full controls.

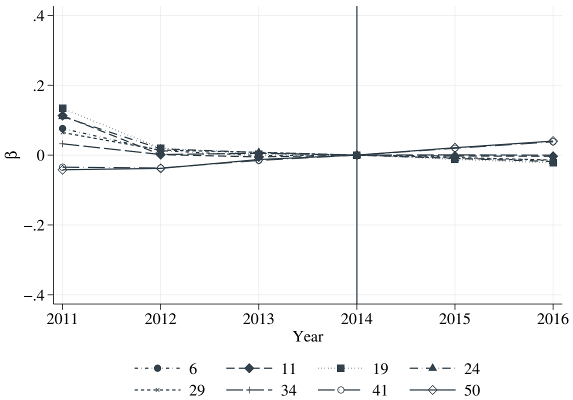


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

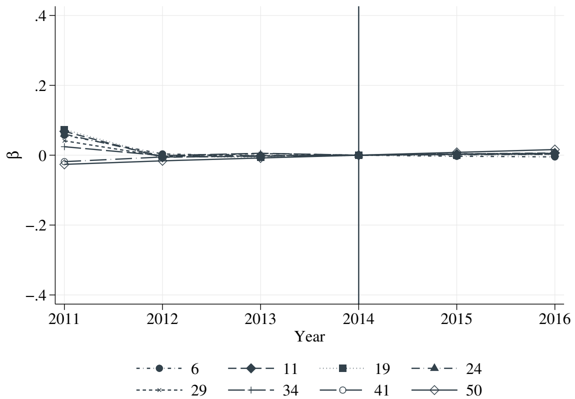
Figure A.20: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: $\leq$ 12 euros/hour. Full controls.



(a) Bite 1: Regions



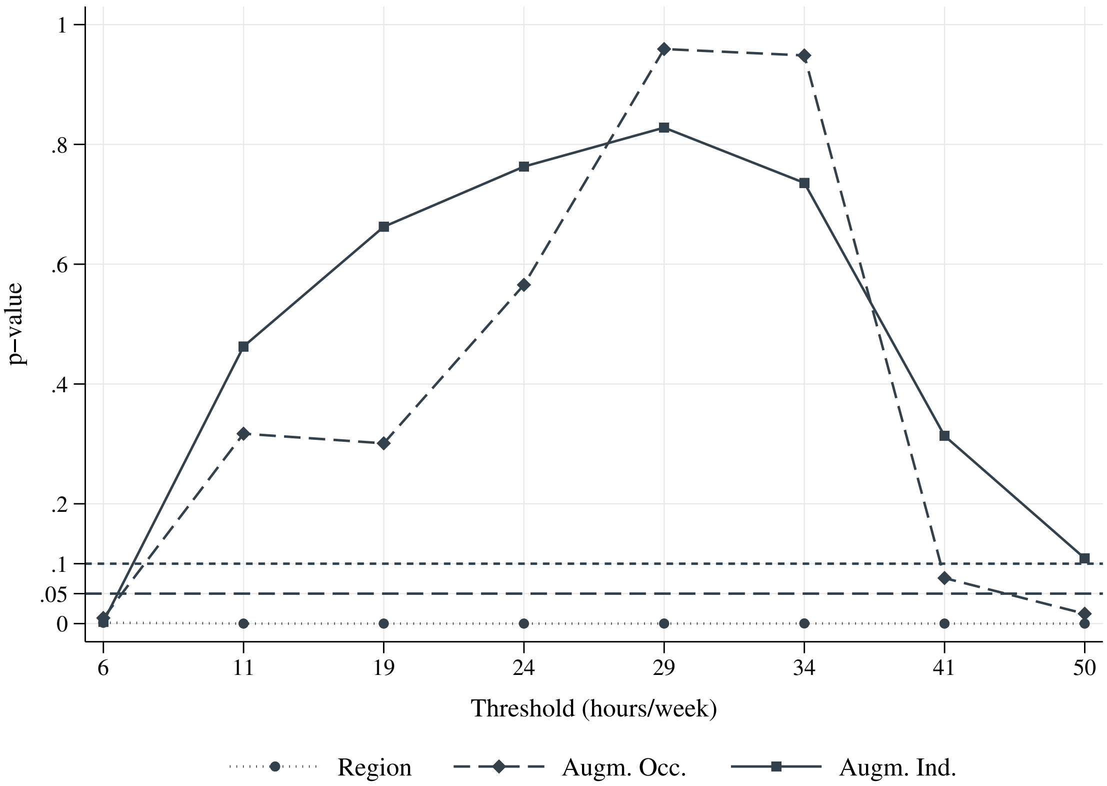
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.21: P-values of joint significance
Weekly hours worked specifications (worker group: $\leq$ 12 euros/hour). Full controls.

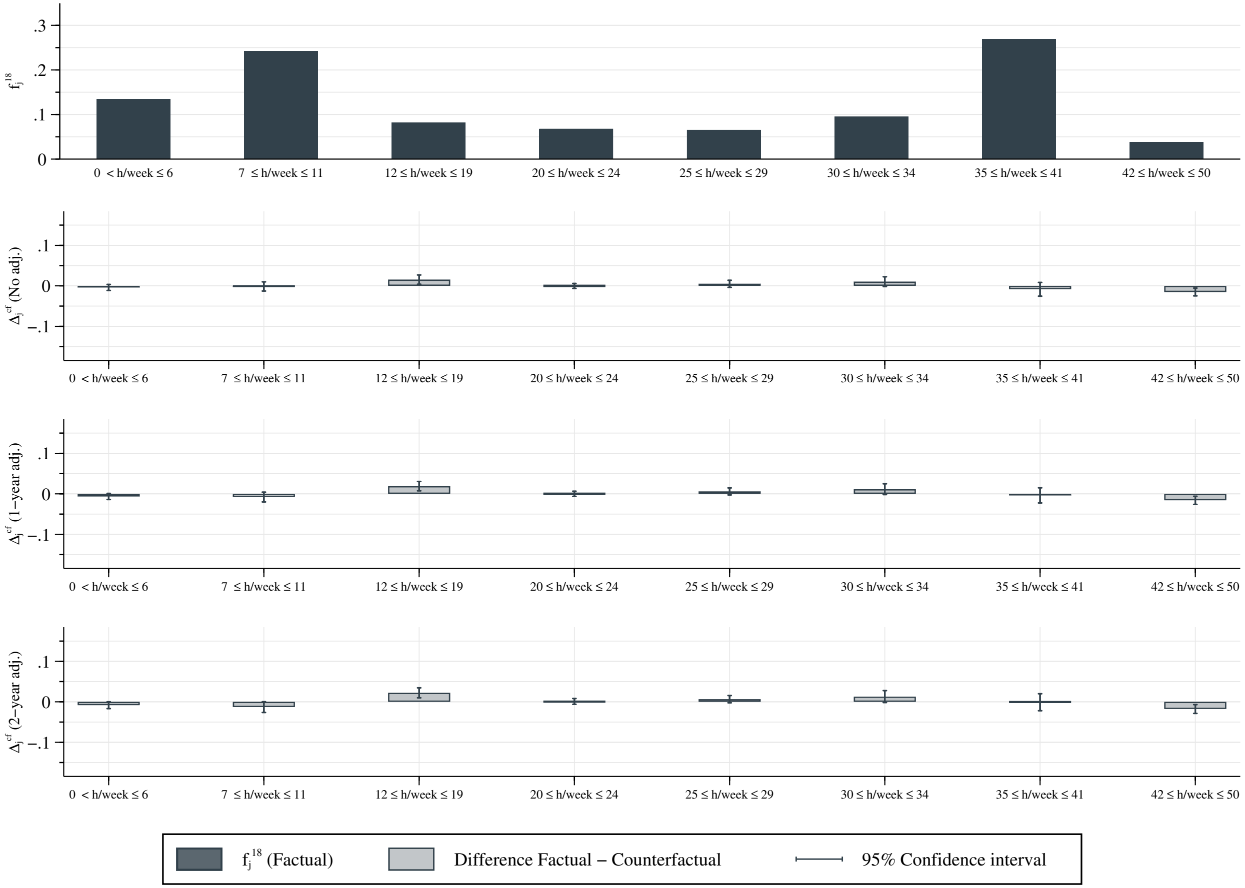


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0:\ \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No firm controls

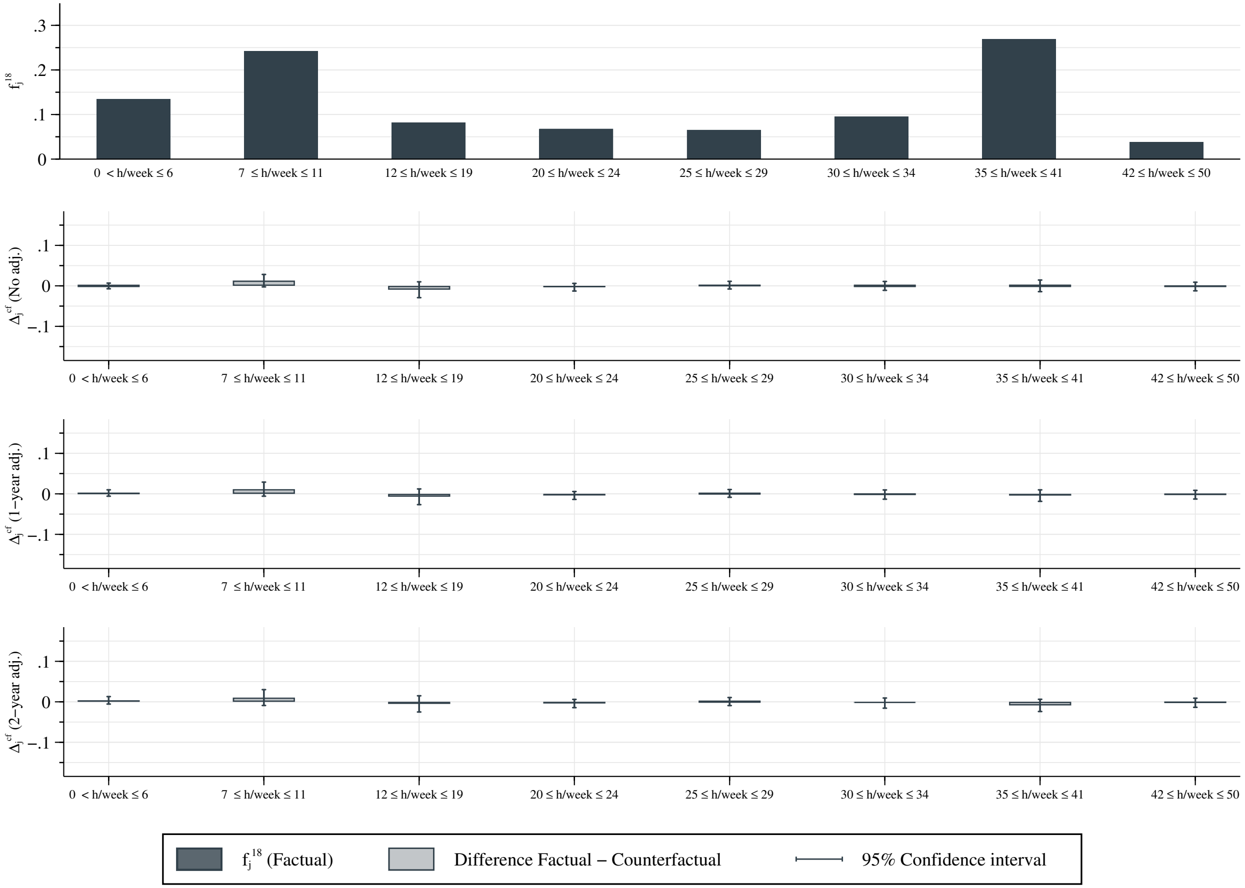
— Intentionally left blank —

Figure A.22: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages ≤ 12 euros/hour. Bite 1: Regions. All trend adjustment regimes. No firm controls.



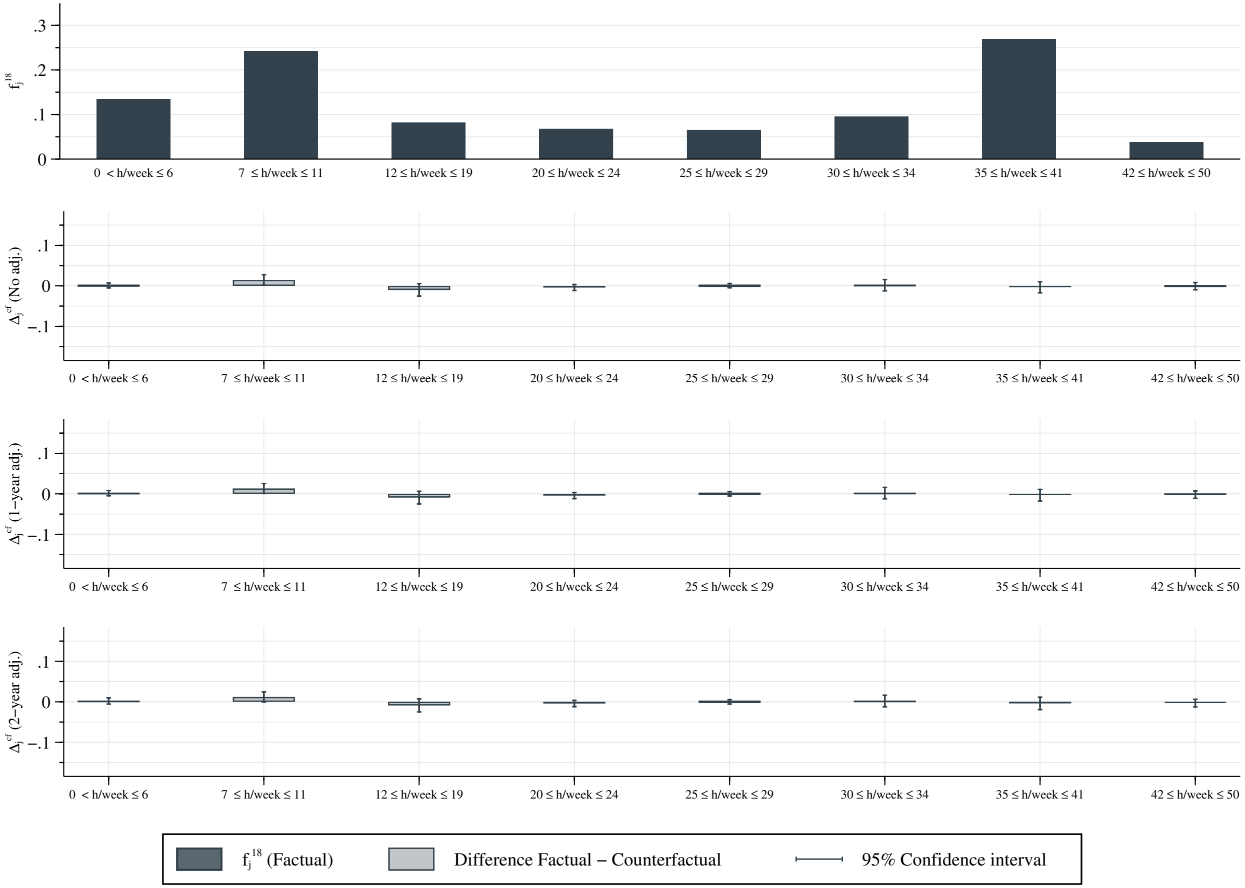
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.23: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages ≤ 12 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. No firm controls.



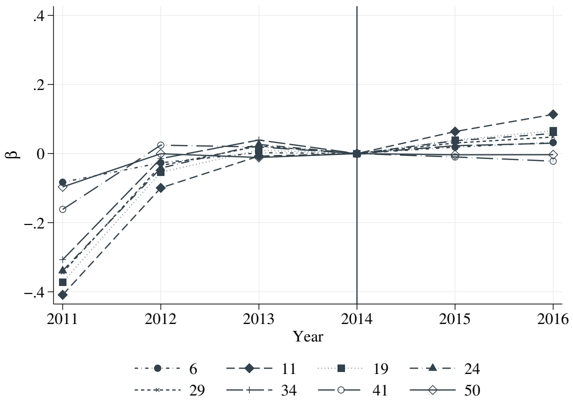
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.24: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. No firm controls.

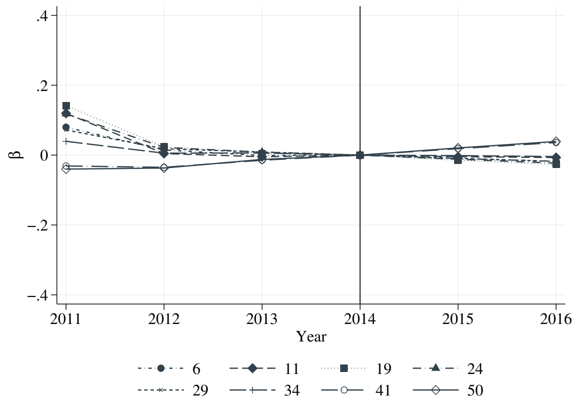


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

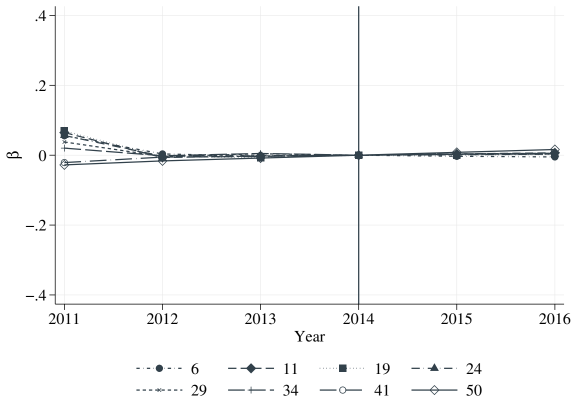
Figure A.25: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: $\leq$ 12 euros/hour. No firm controls.



(a) Bite 1: Regions



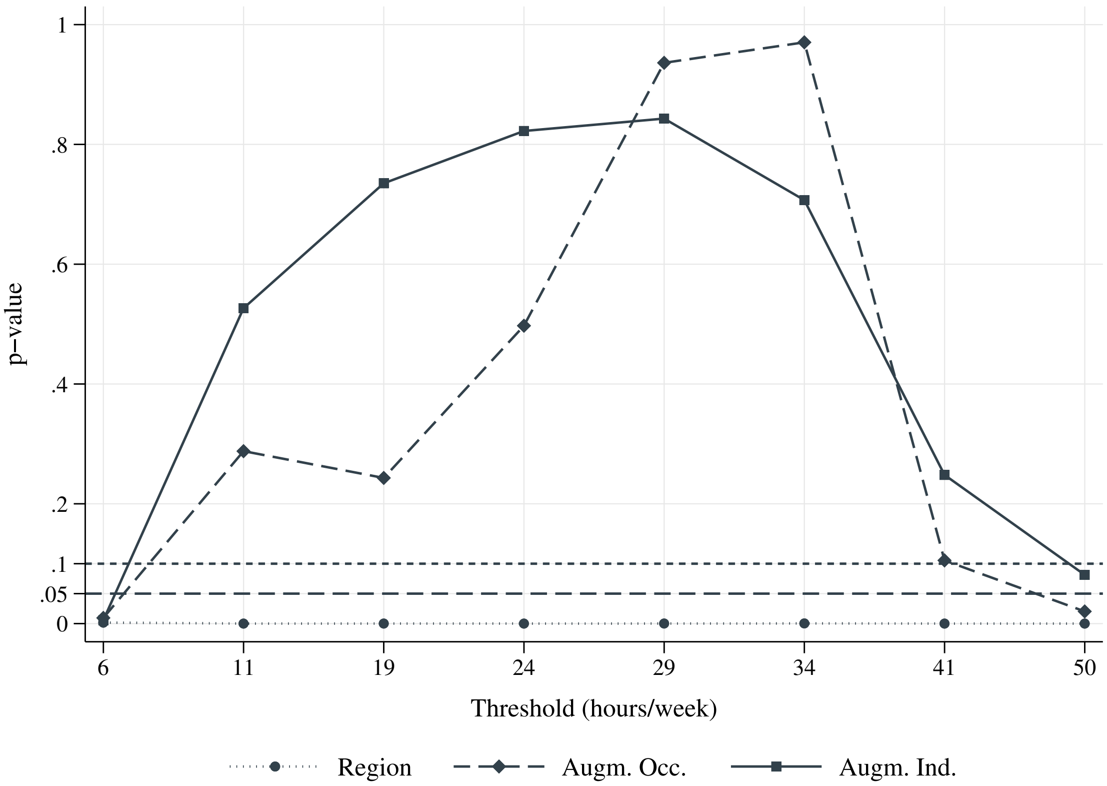
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.26: P-values of joint significance
Weekly hours worked specifications (worker group: $\leq$ 12 euros/hour). No firm controls.

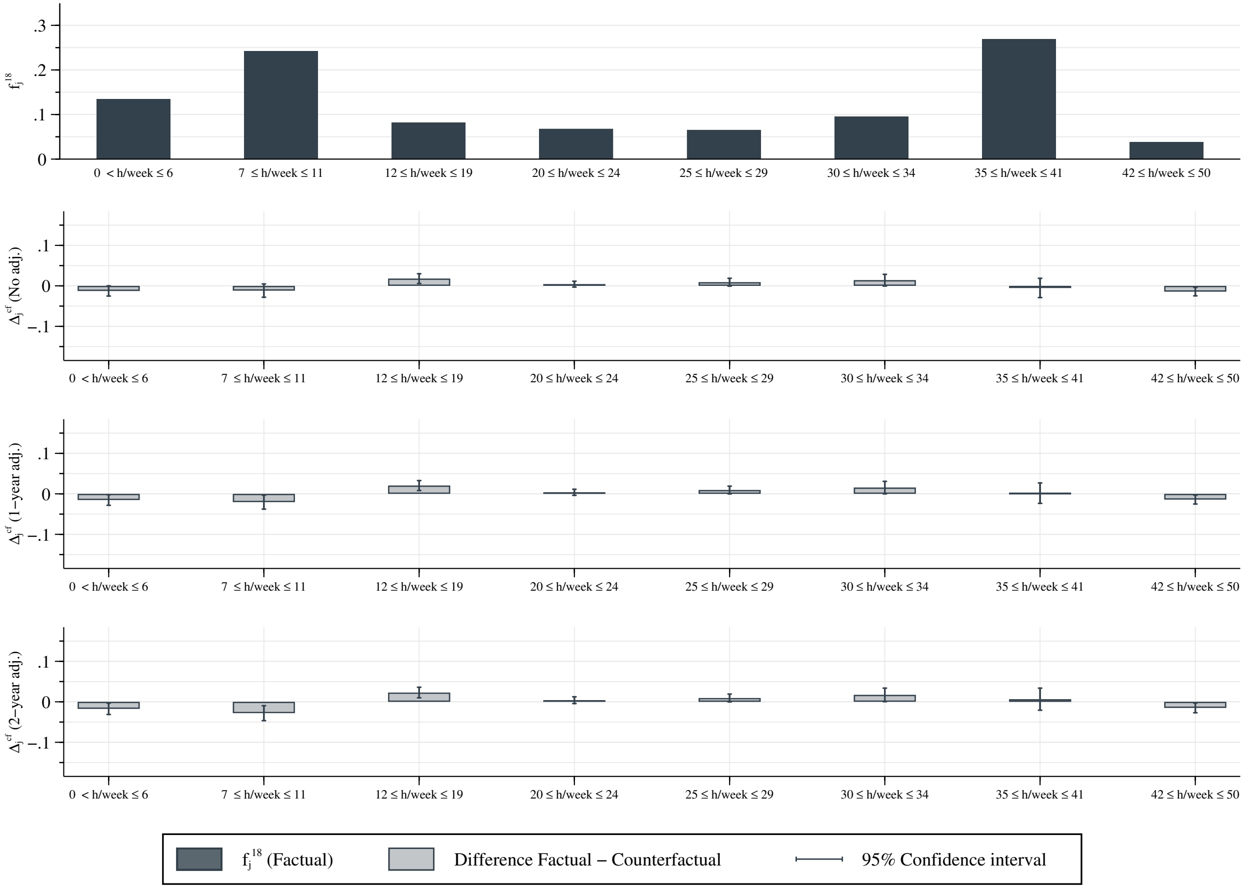


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates H_0 : $\beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No controls

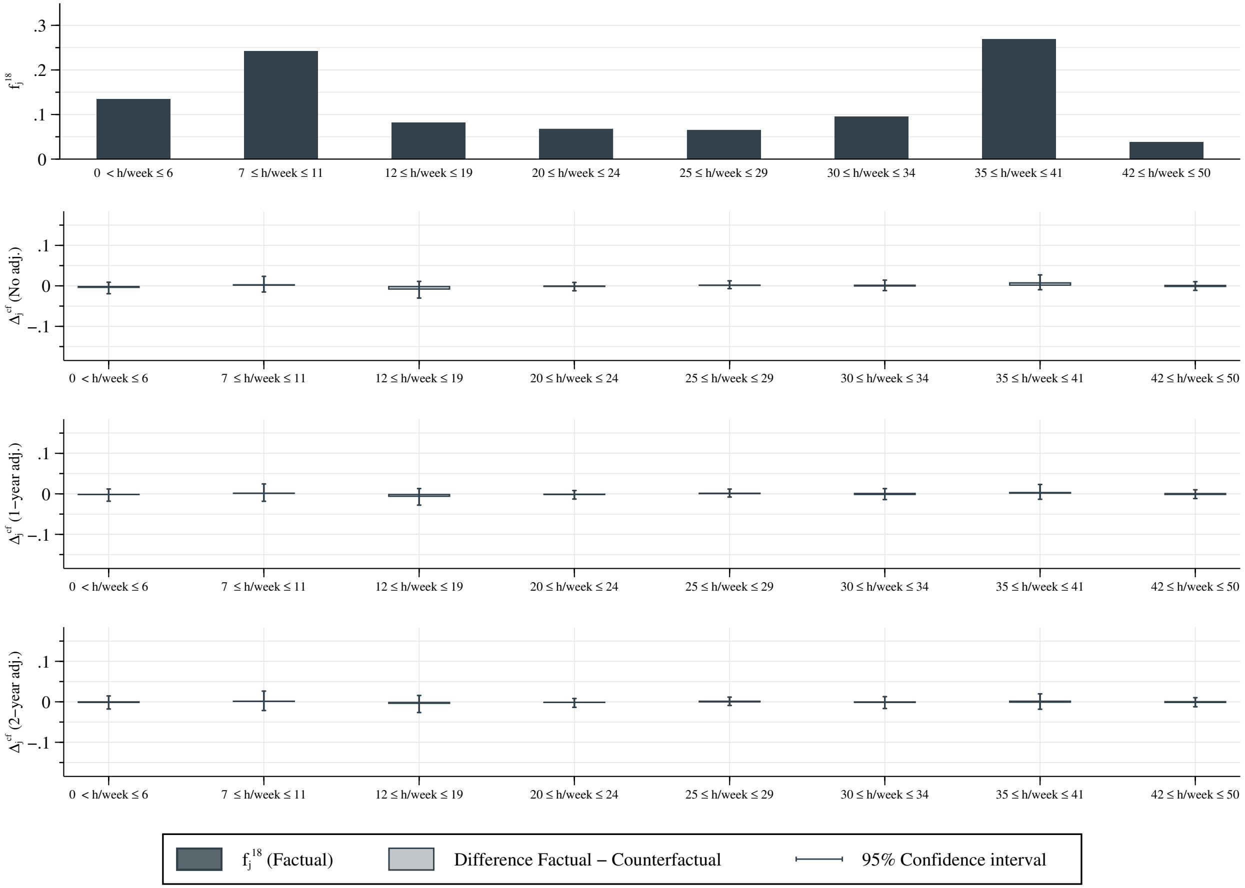
— Intentionally left blank —

Figure A.27: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 1: Regions. All trend adjustment regimes. No controls.



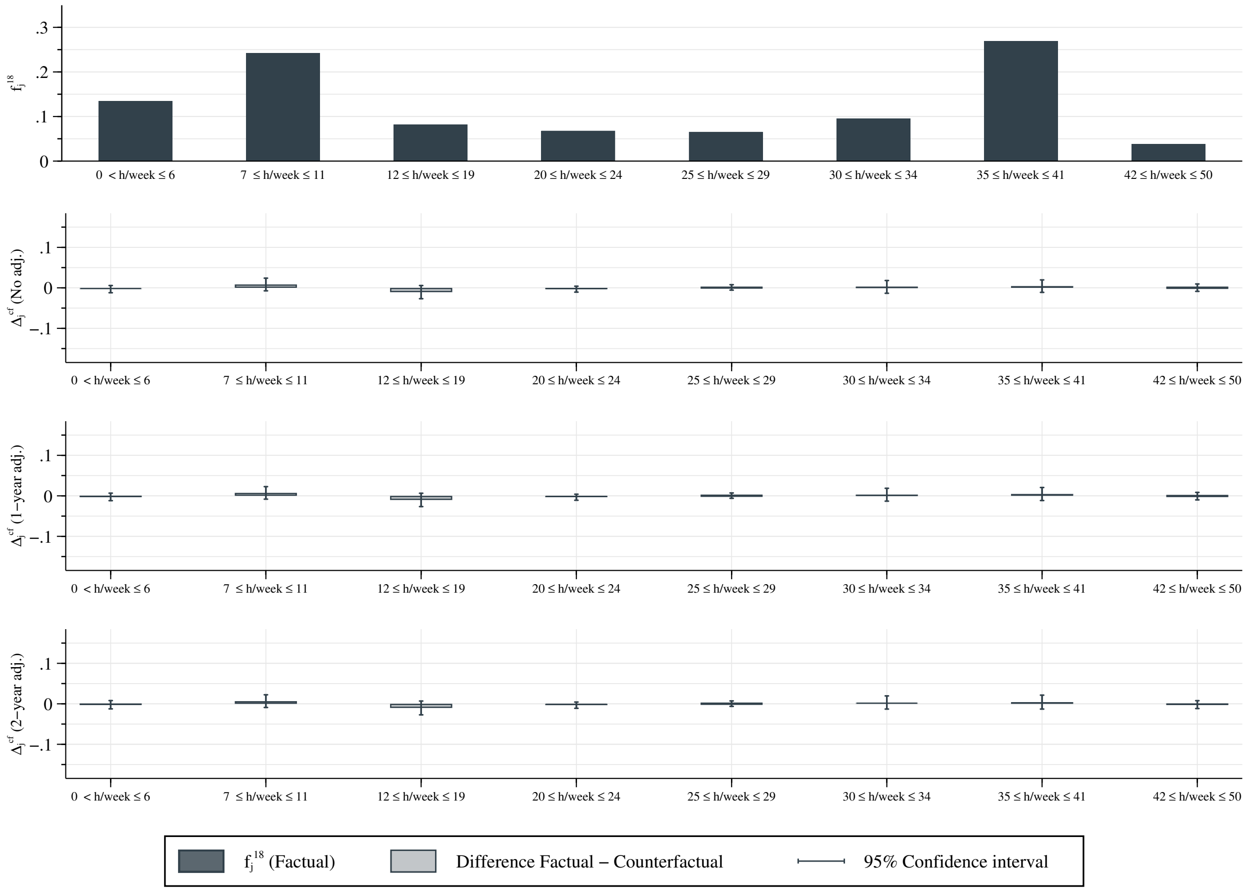
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.28: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. No controls.



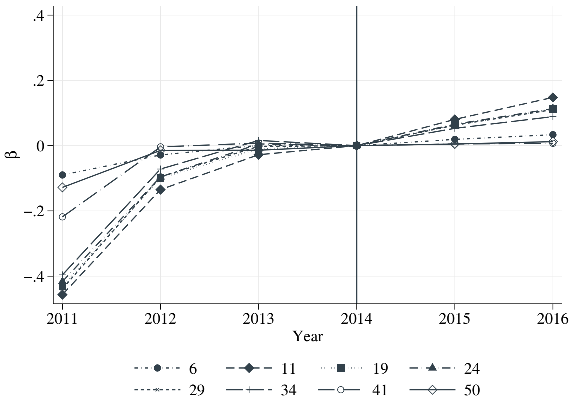
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.29: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages $\leq$ 12 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. No controls.

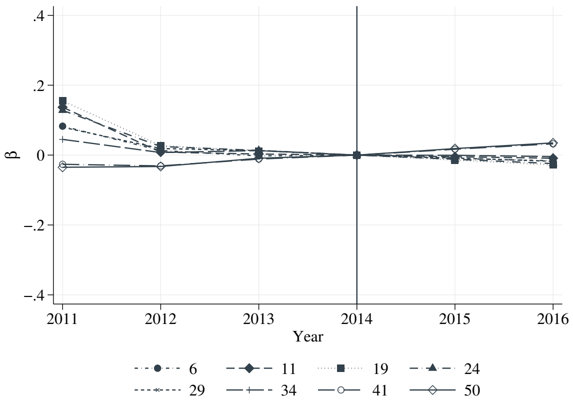


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

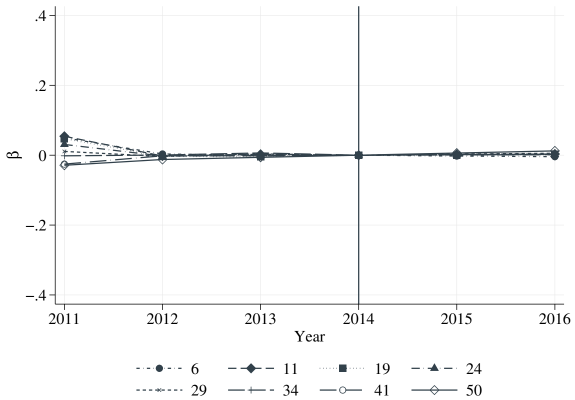
Figure A.30: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: $\leq$ 12 euros/hour. No controls.



(a) Bite 1: Regions



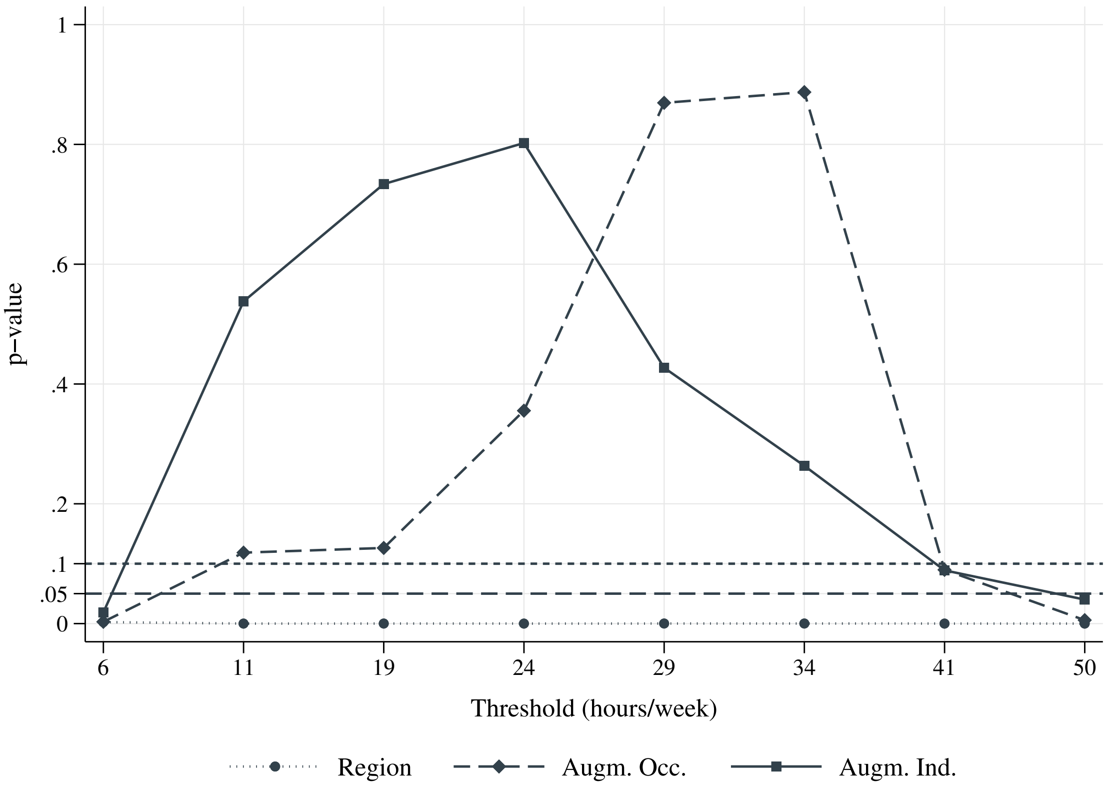
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.31: P-values of joint significance
Weekly hours worked specifications (worker group: $\leq$ 12 euros/hour). No controls.

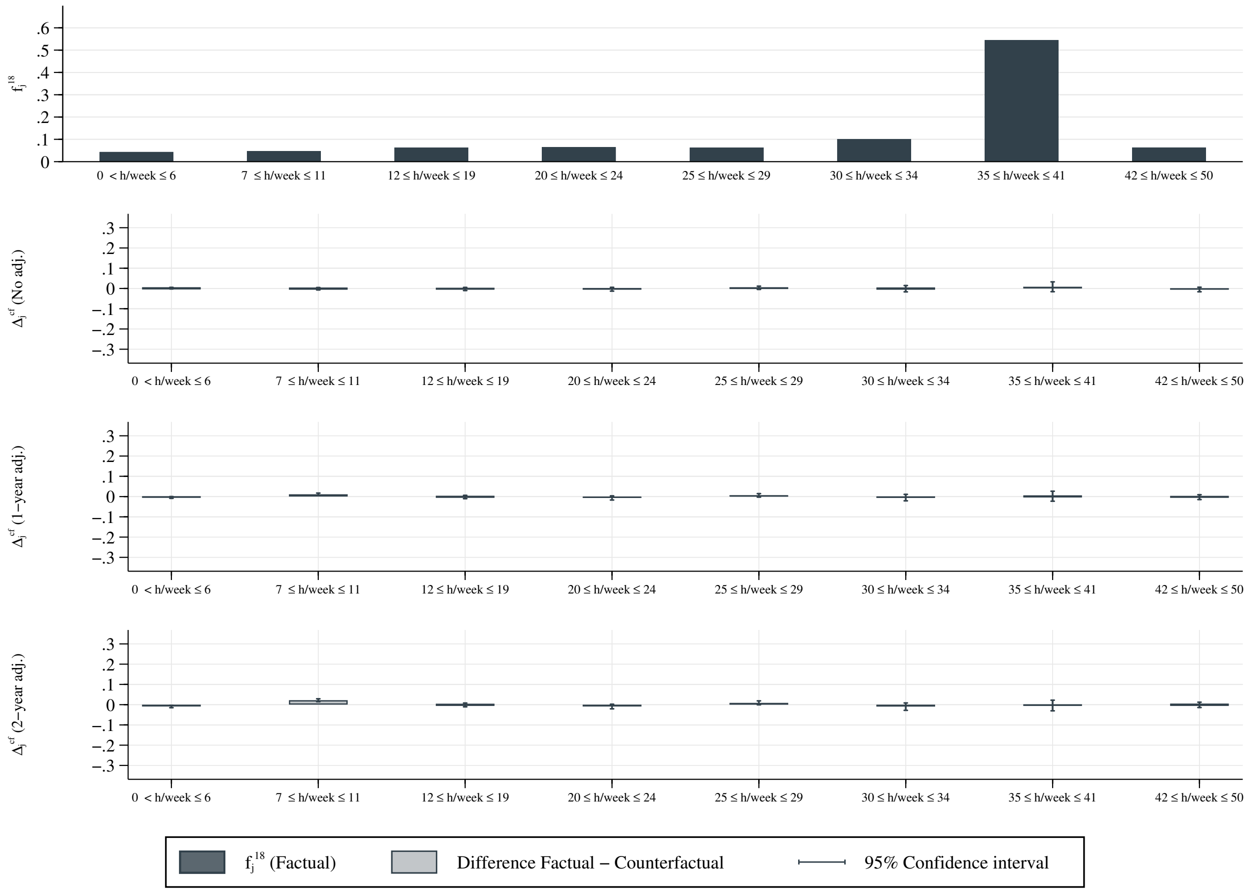


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

Hourly wages between 12 and 16 euros/hour

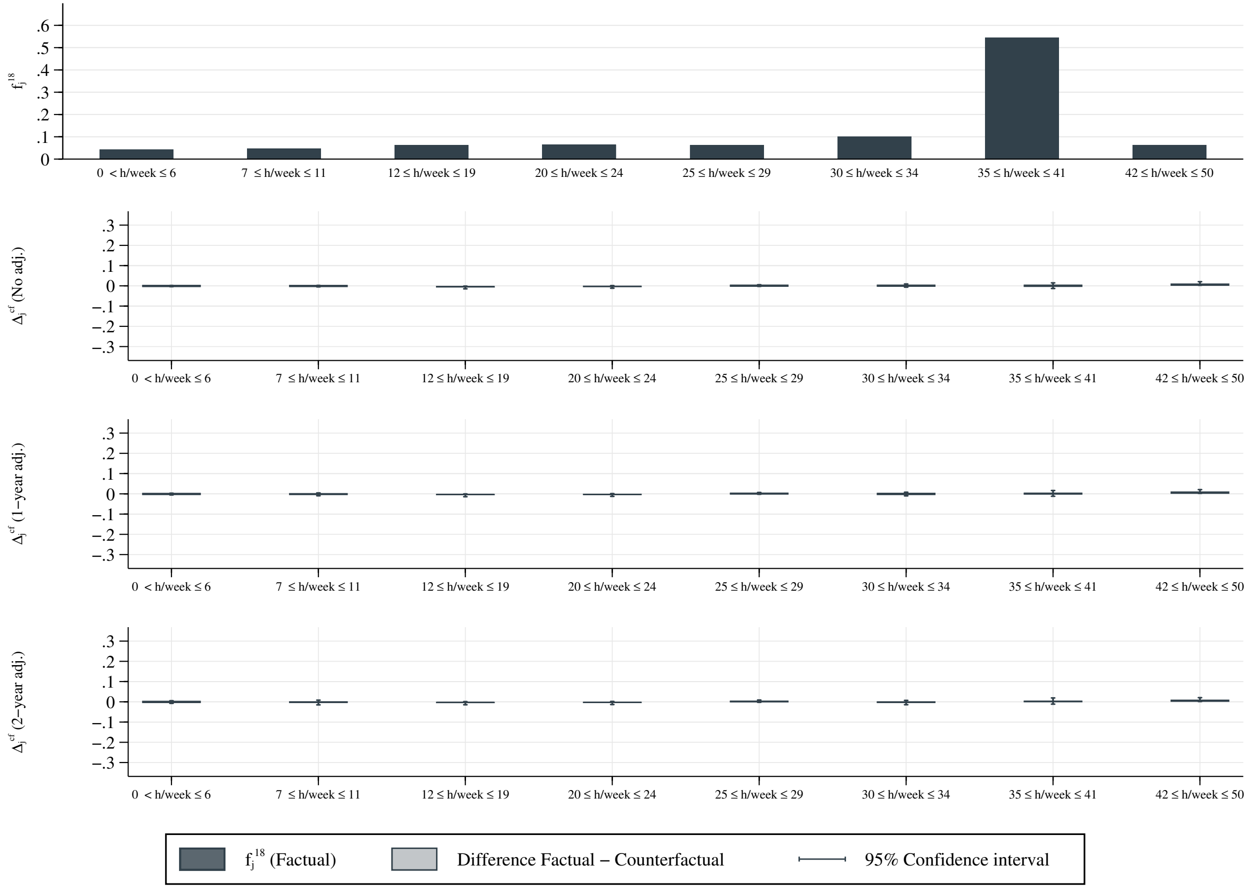
— Intentionally left blank —

Figure A.32: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. Full controls.



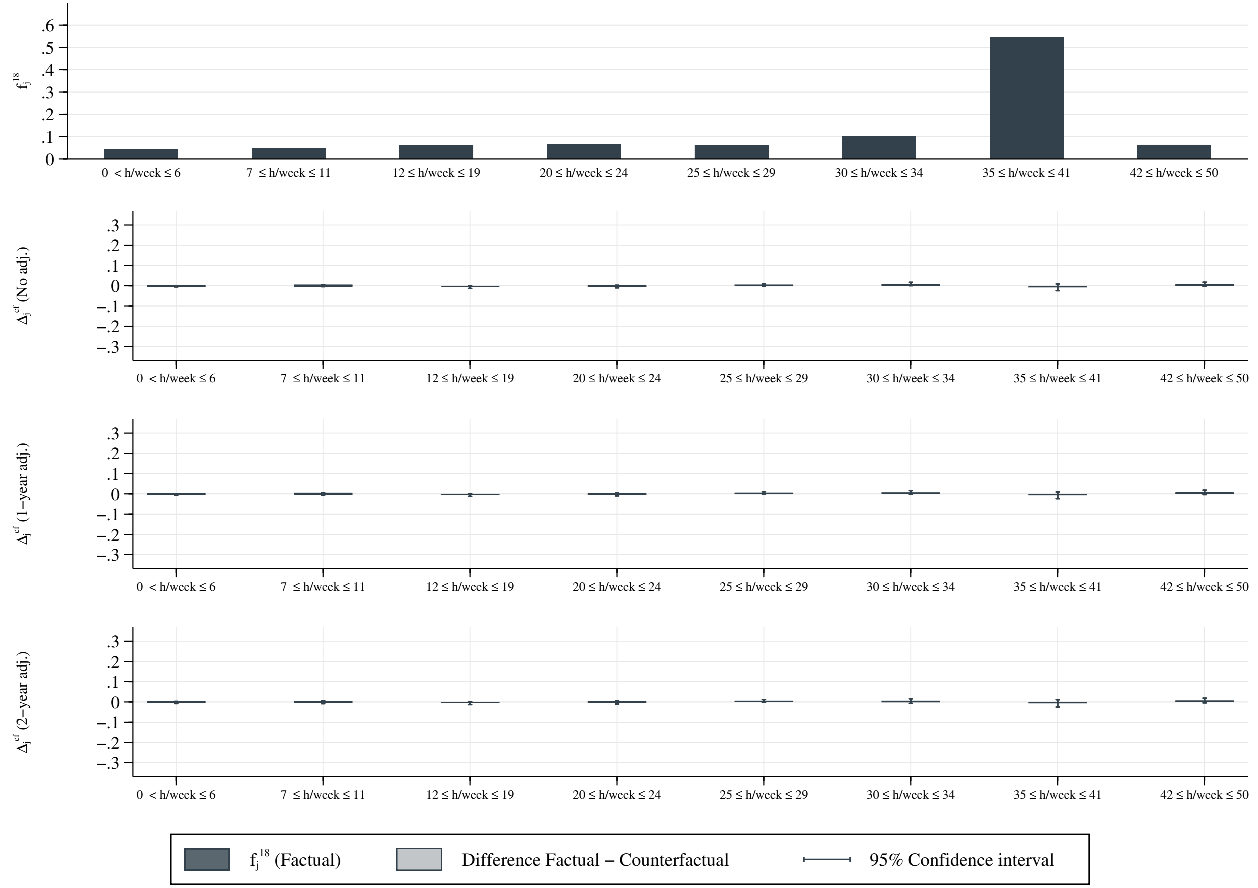
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.33: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. Full controls.



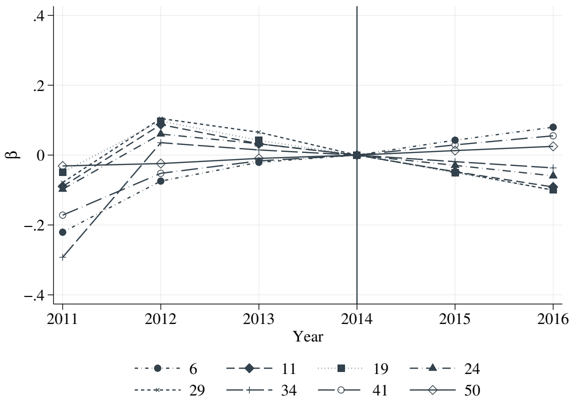
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.34: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. Full controls.

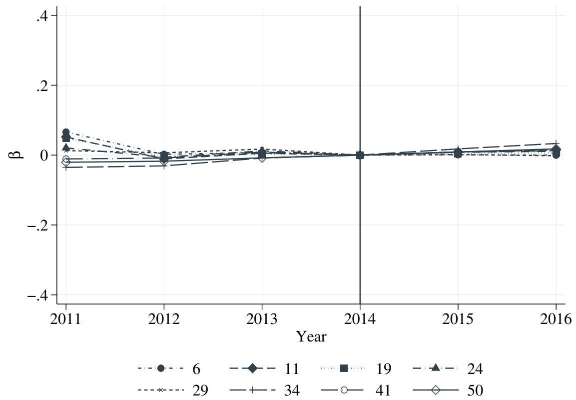


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

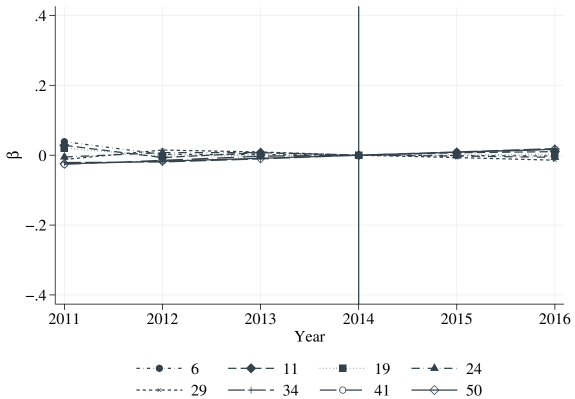
Figure A.35: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Weekly working hours, worker group: hourly wages between 12 and 16 euros/hour. Full controls.



(a) Bite 1: Regions



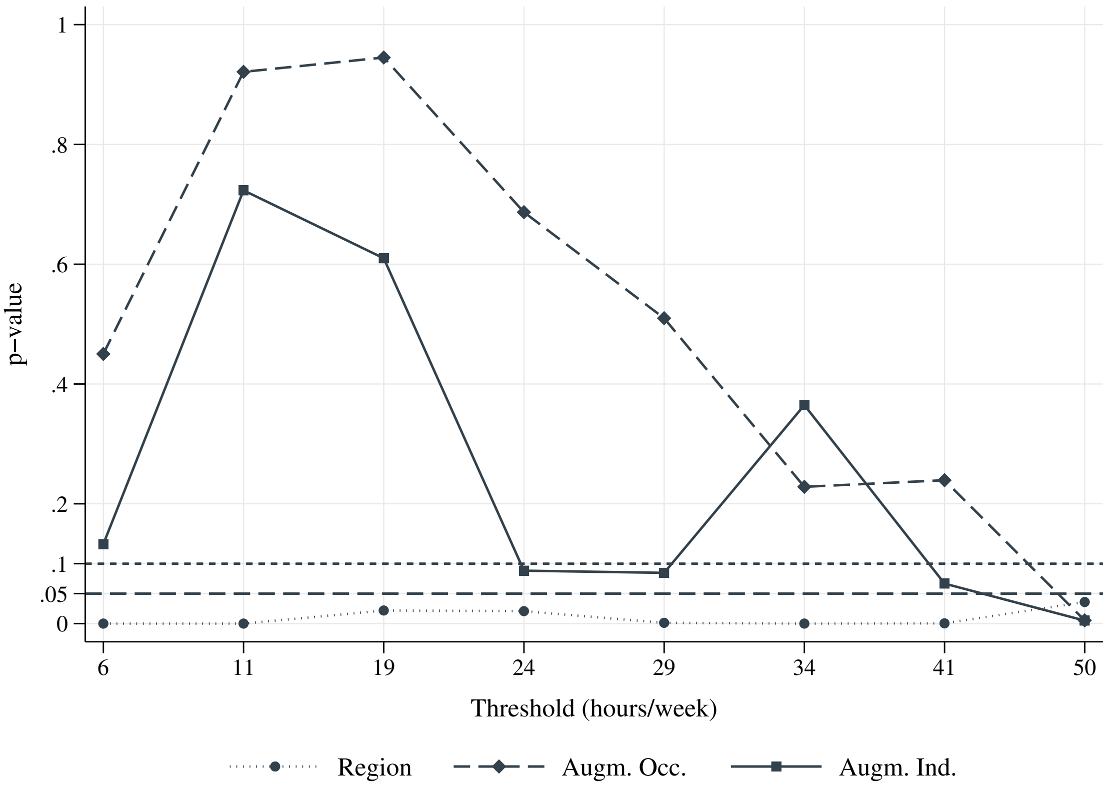
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.36: P-values of joint significance. Weekly hours worked specifications. Worker group: hourly wage between 12 and 16 euros/hour. Full controls.

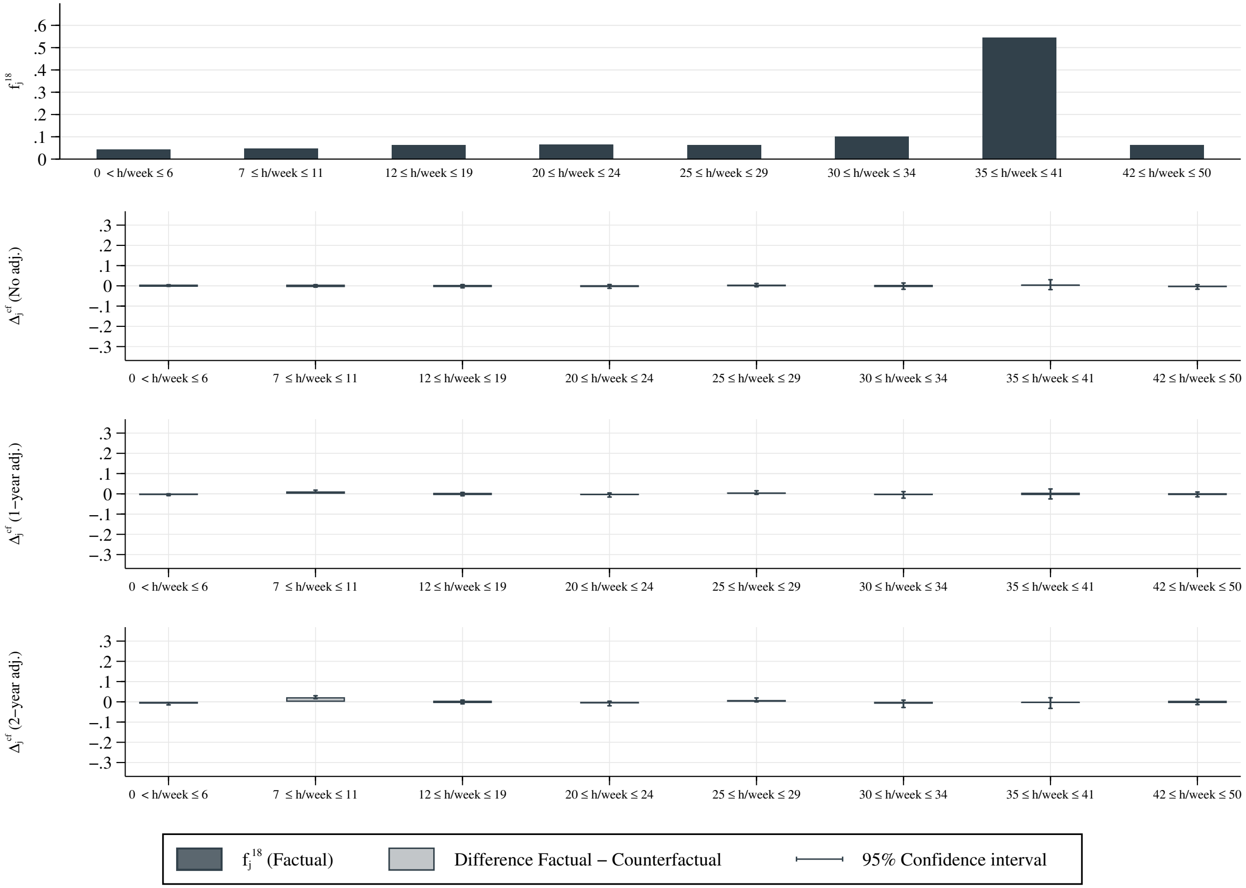


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No firm controls

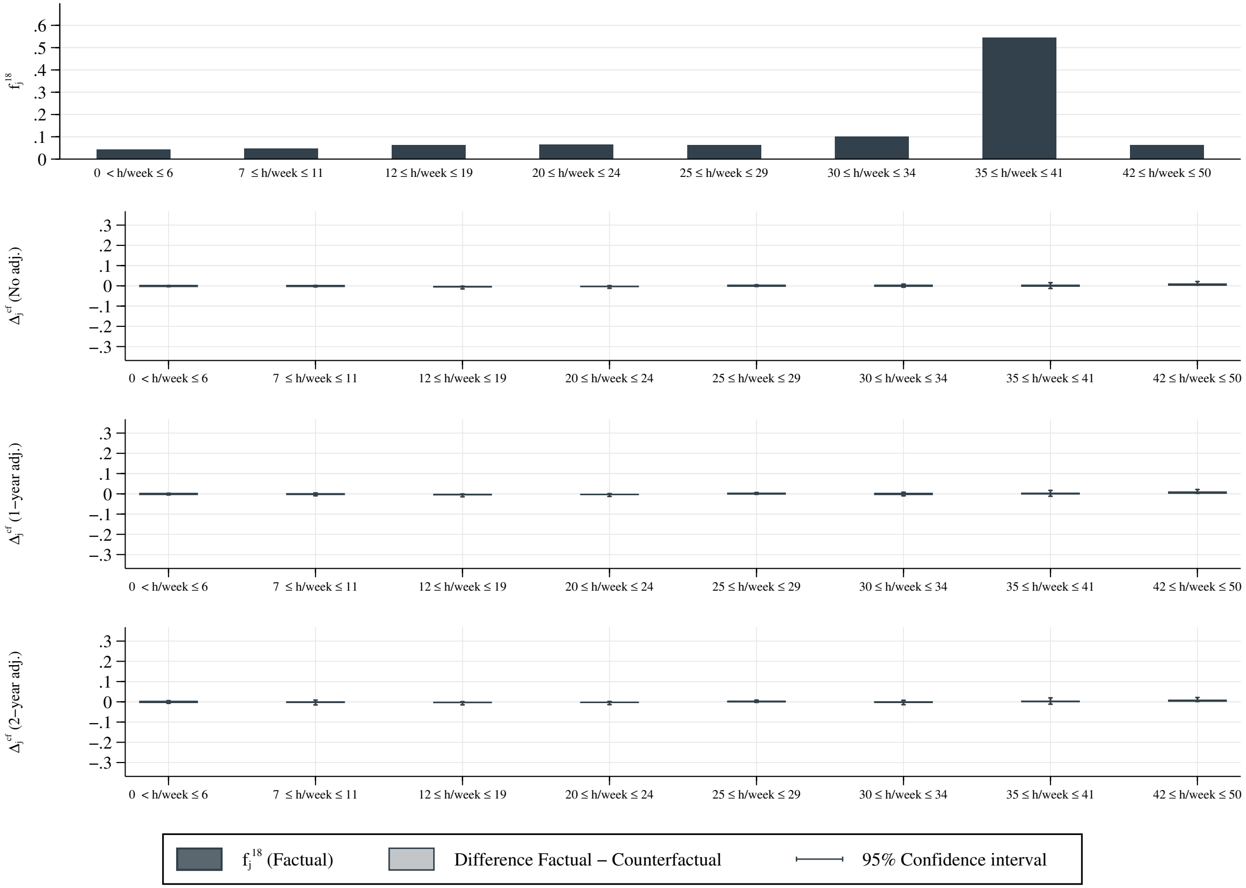
— Intentionally left blank —

Figure A.37: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. No firm controls.



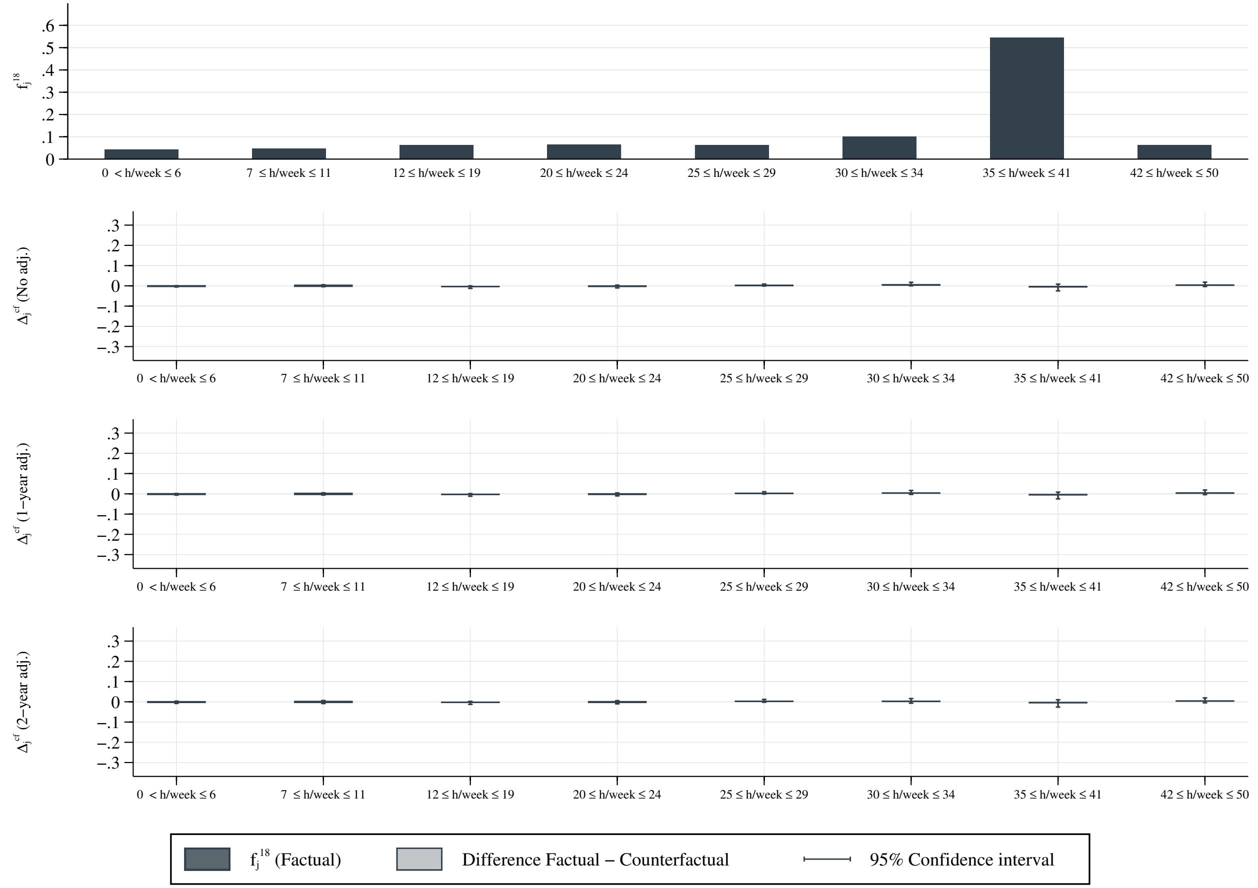
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.38: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 2: Augm. occupations. All trend adjustment regimes. No firm controls.



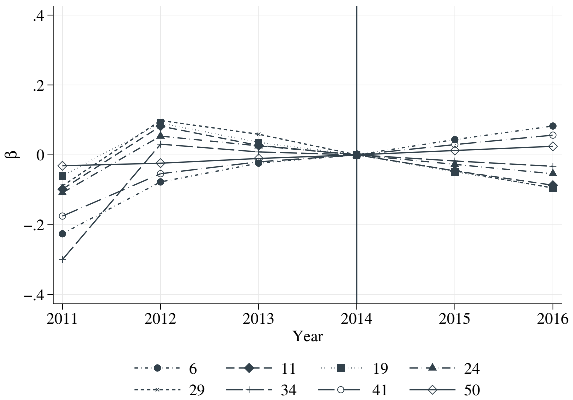
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.39: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 3: Augm. industries. All trend adjustment regimes. No firm controls.

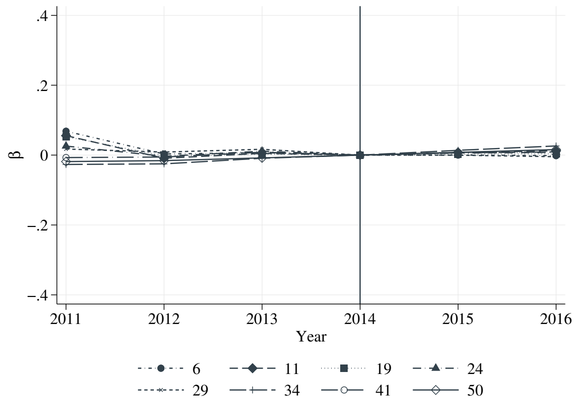


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

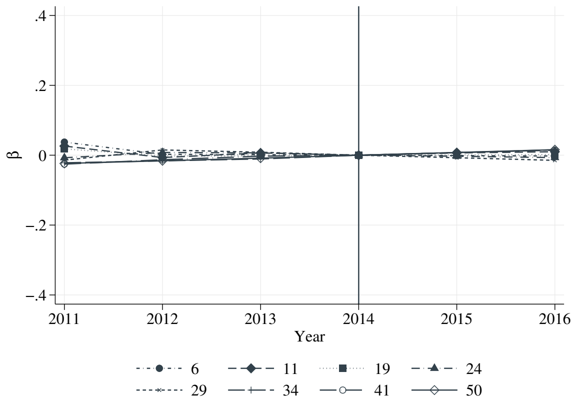
Figure A.40: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Weekly working hours, worker group: hourly wages between 12 and 16 euros/hour. No firm controls.



(a) Bite 1: Regions



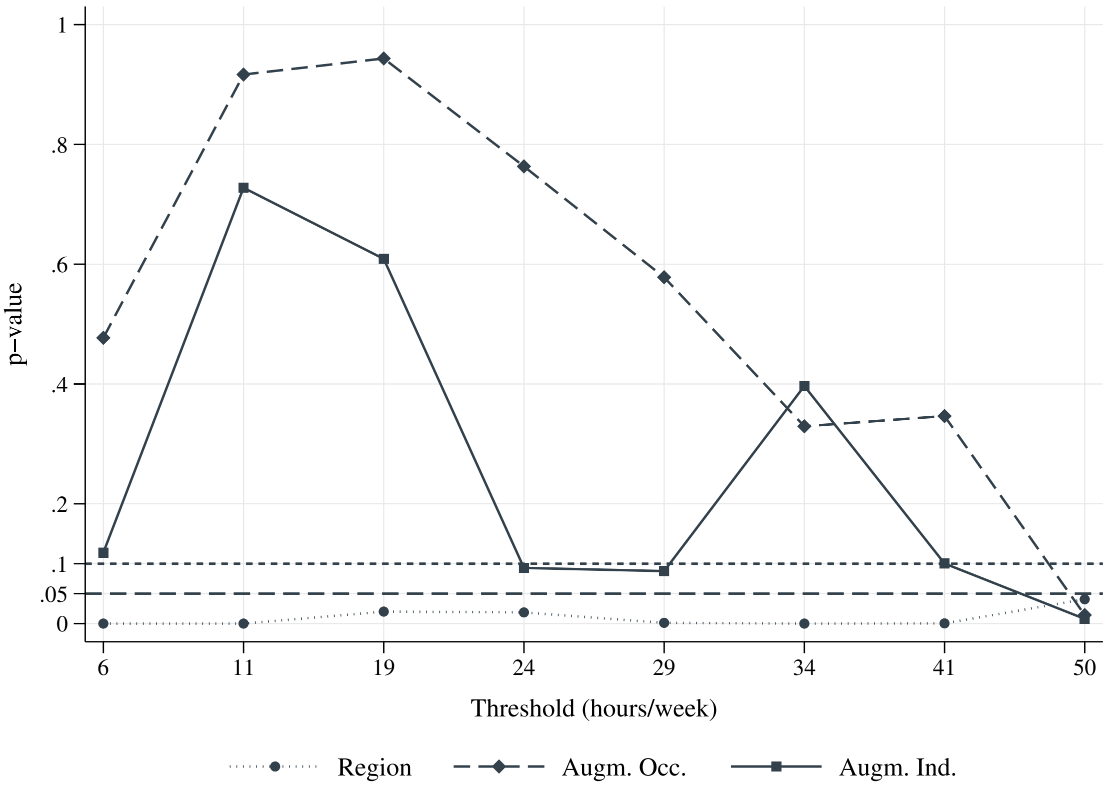
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.41: P-values of joint significance. Weekly hours worked specifications. Worker group: hourly wage between 12 and 16 euros/hour. No firm controls.

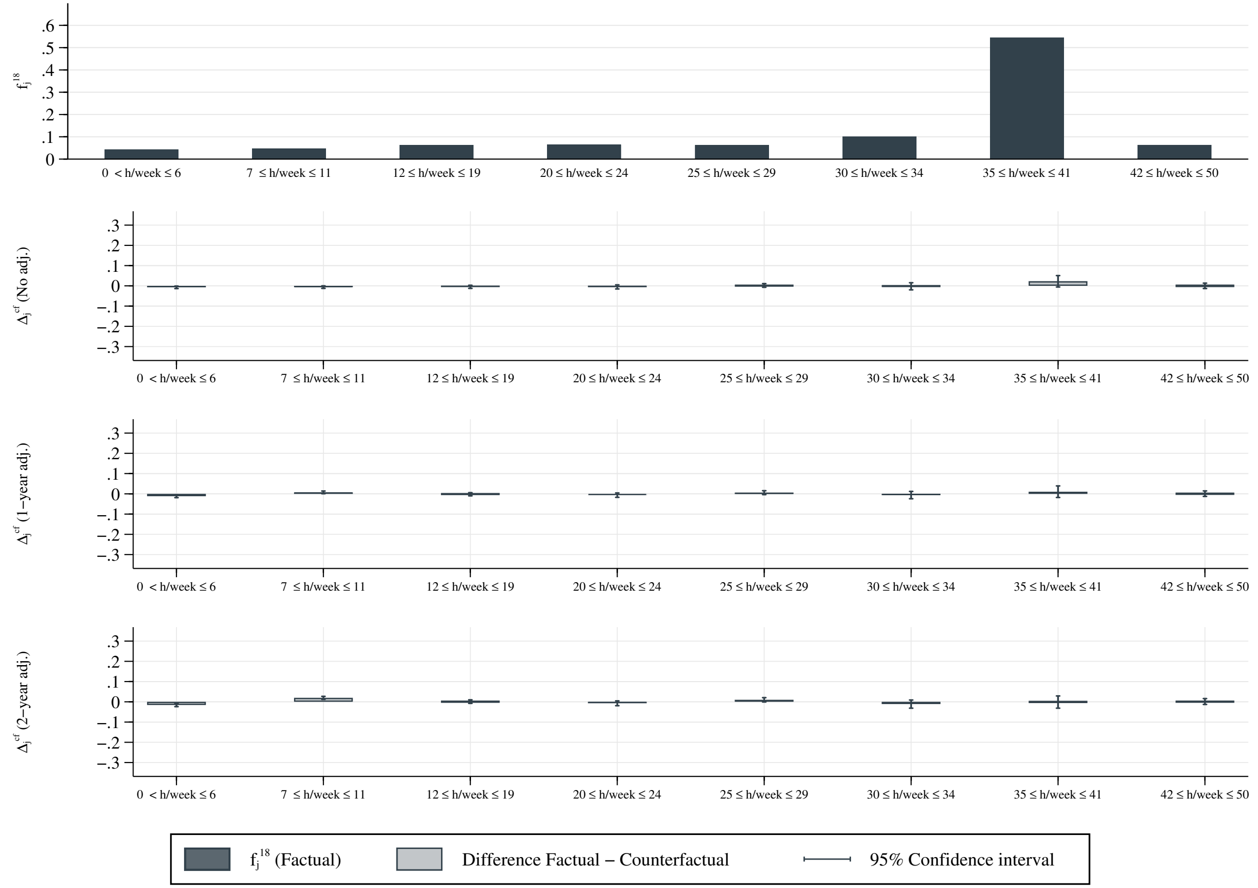


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0:\ \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No controls

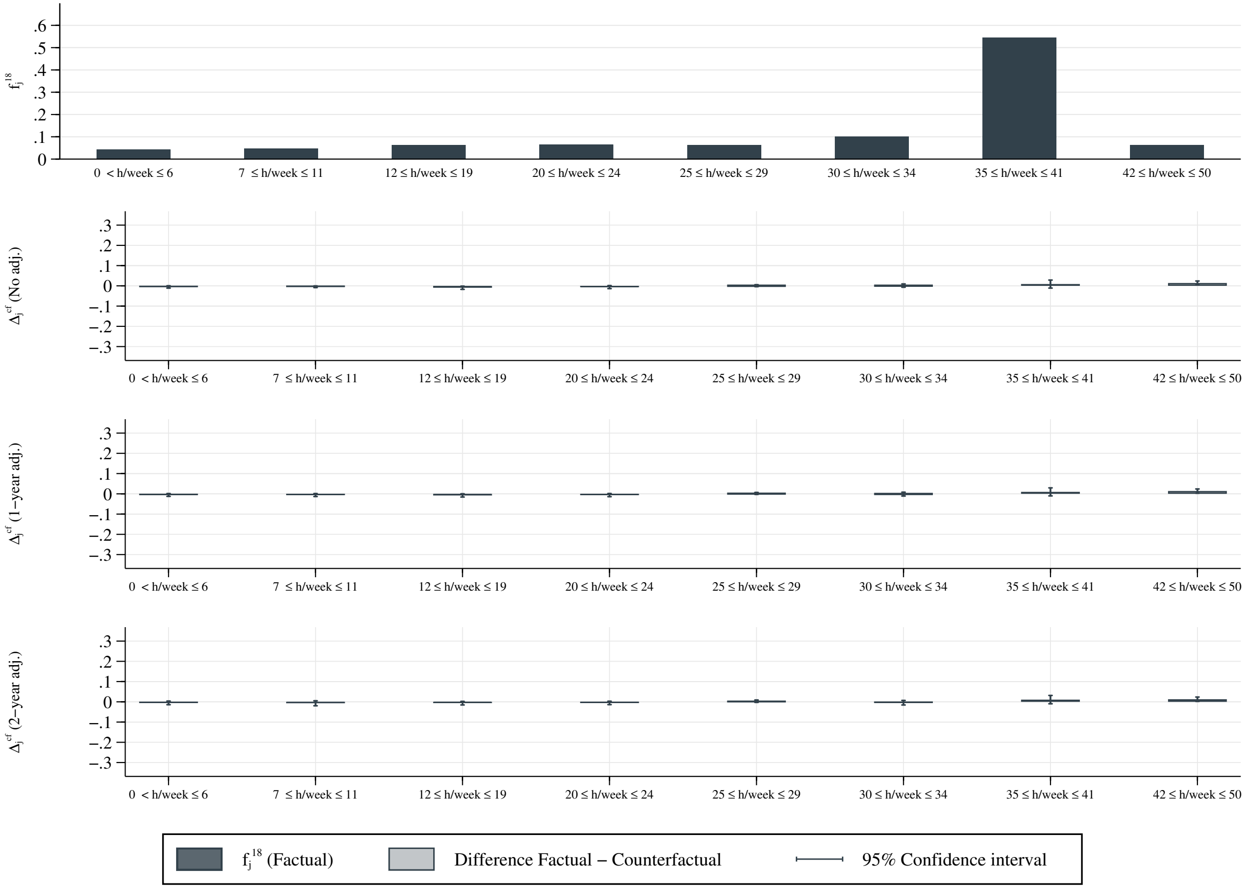
— Intentionally left blank —

Figure A.42: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. No controls.



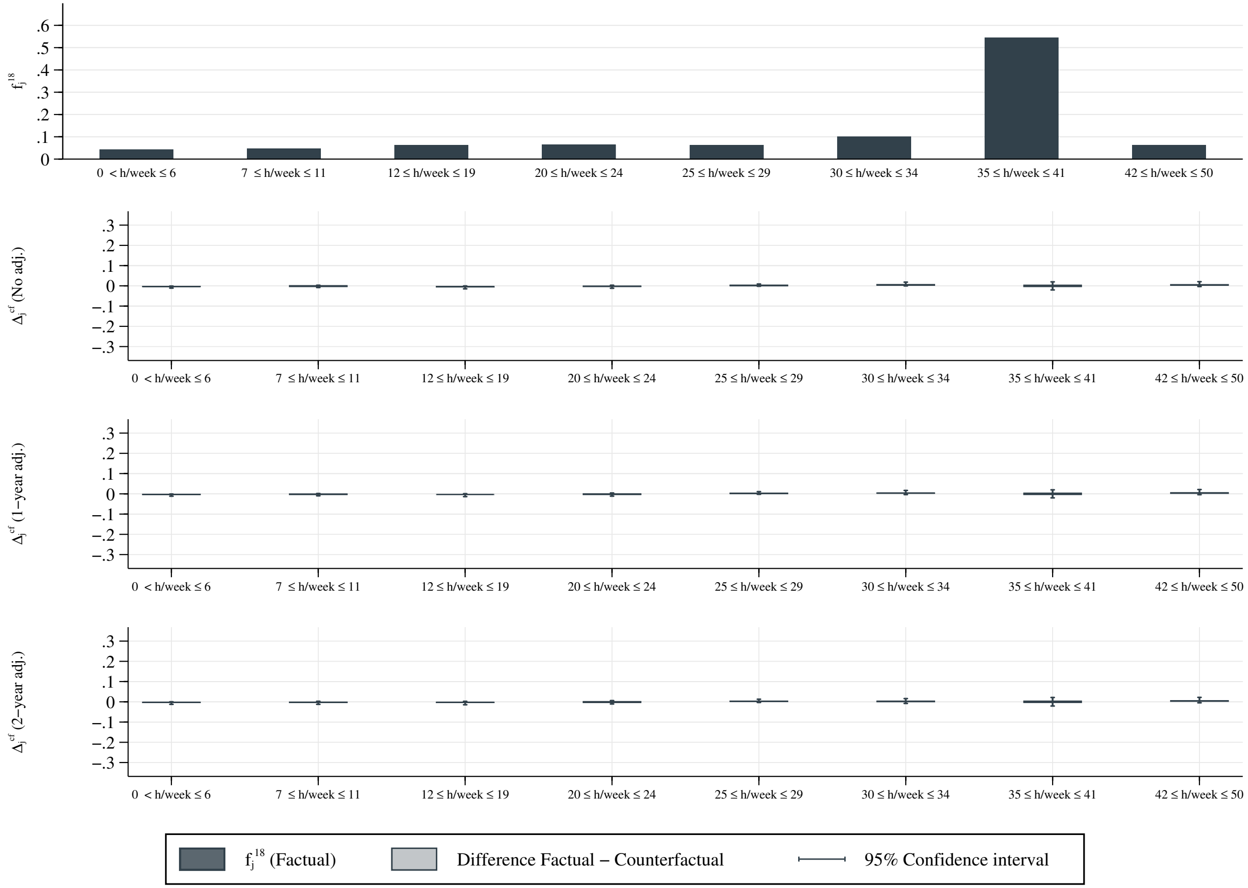
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.43: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour Bite 2: Augmented occupations. All trend adjustment regimes. No controls.



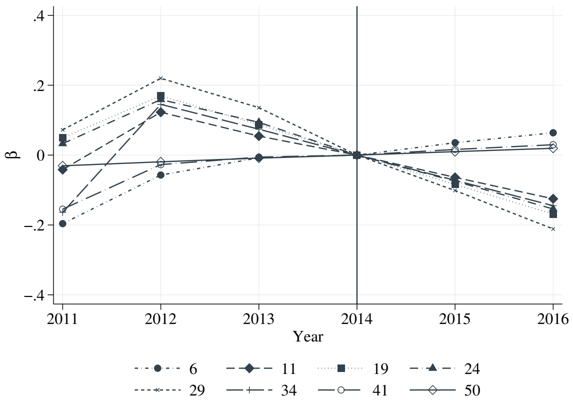
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.44: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages between 12 and 16 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. No controls.

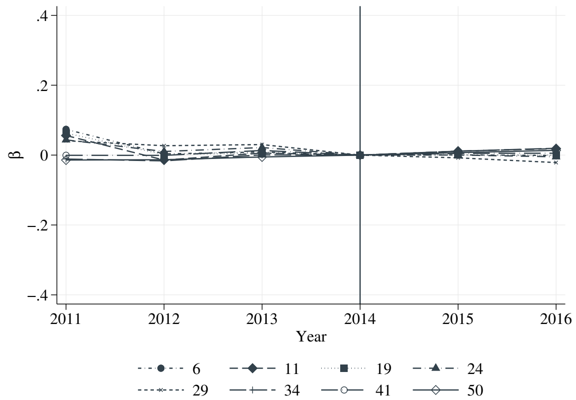


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

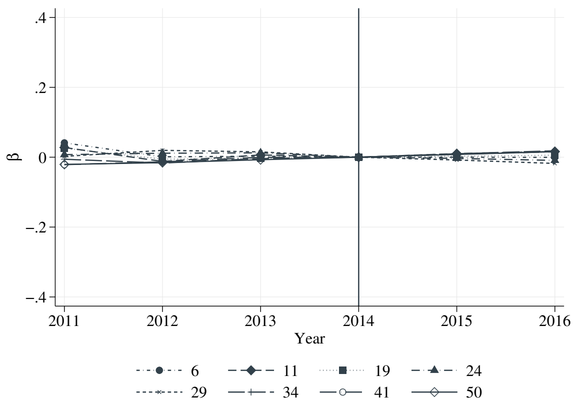
Figure A.45: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Weekly working hours, worker group: hourly wages between 12 and 16 euros/hour. No controls.



(a) Bite 1: Regions



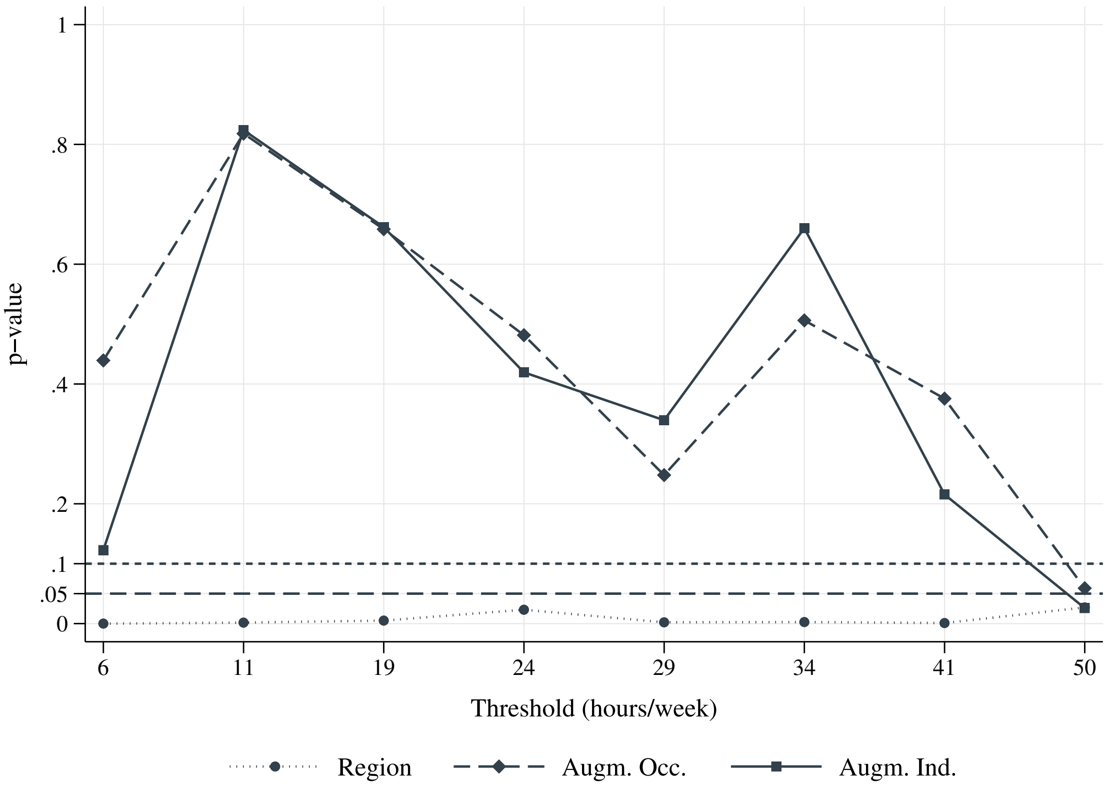
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.46: P-values of joint significance. Weekly hours worked specifications. No controls.

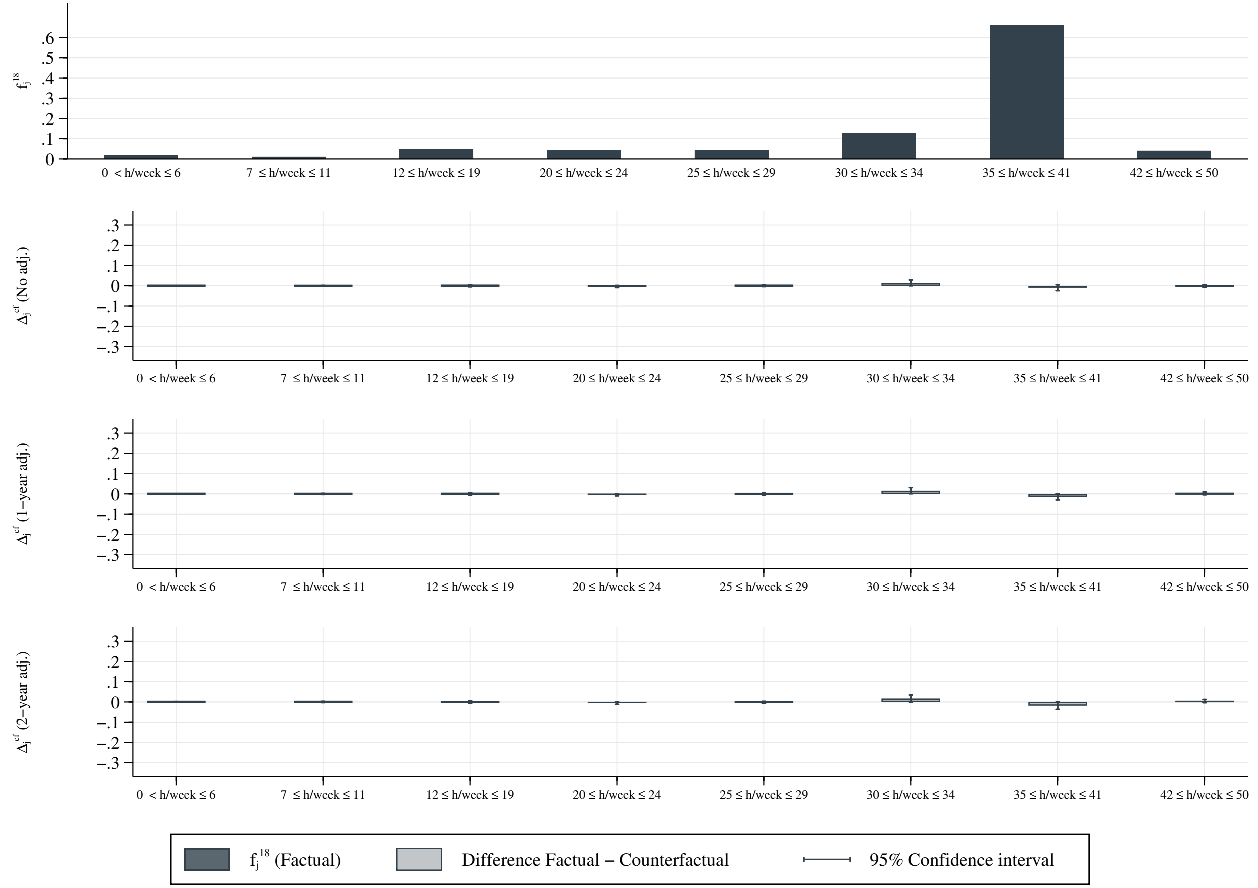


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

Hourly wages above 16 euros/hour

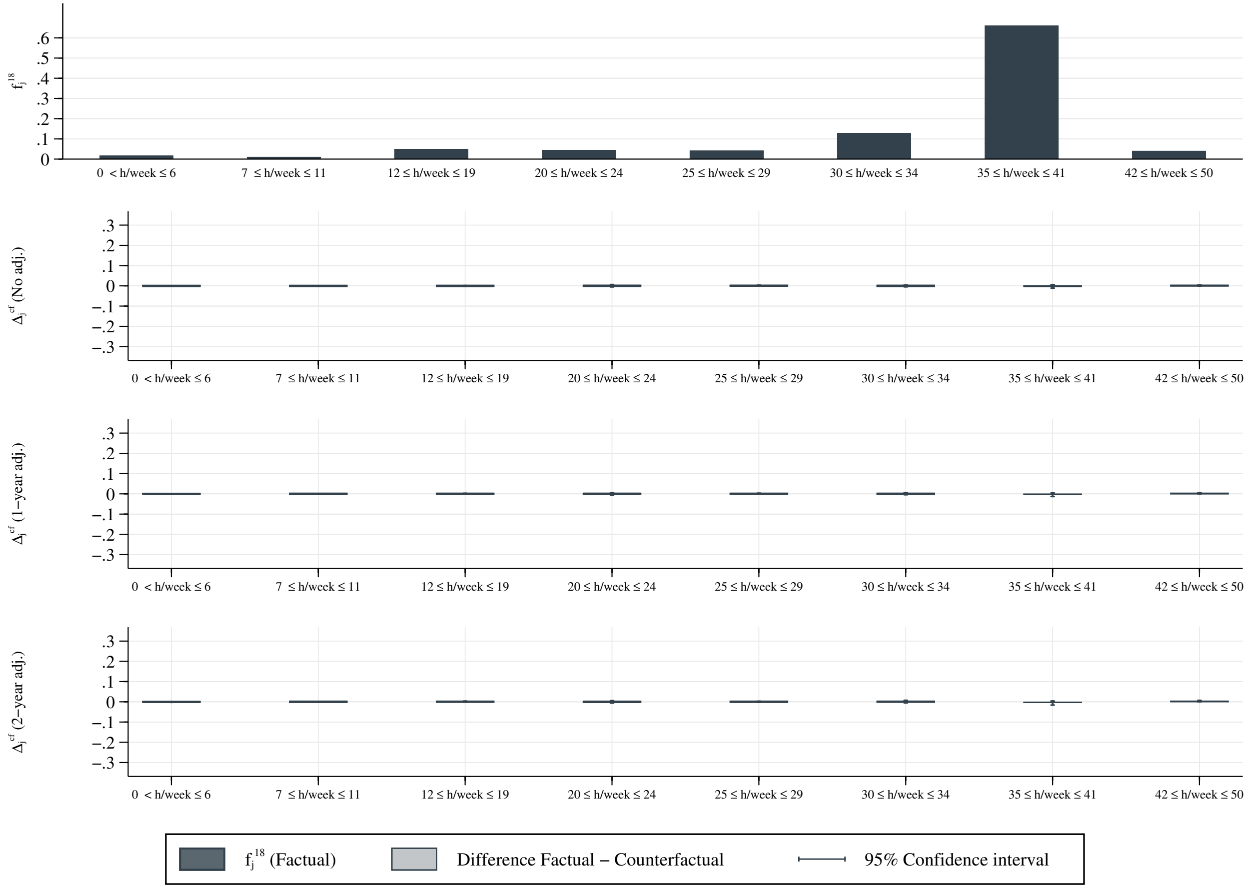
— Intentionally left blank —

Figure A.47: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. Full controls.



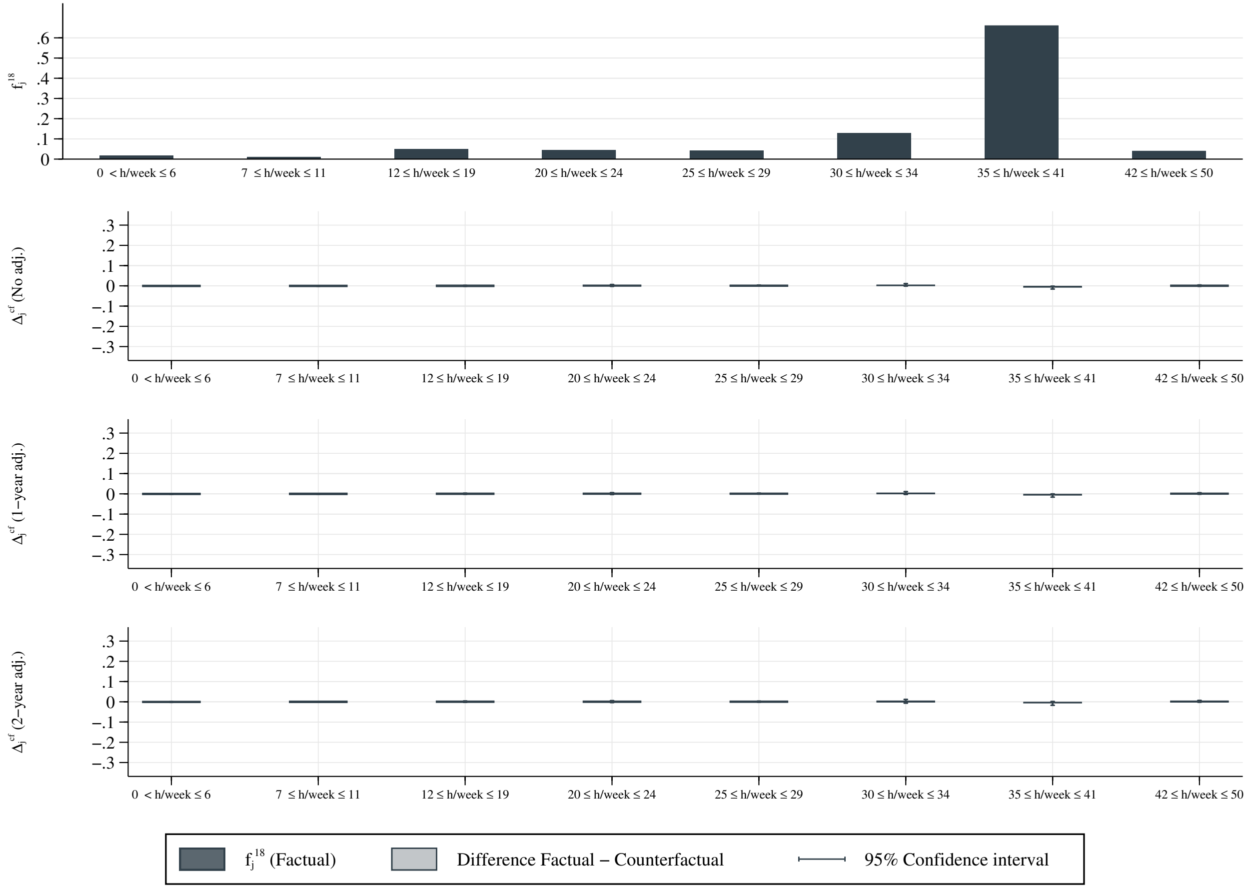
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.48: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. Full controls.



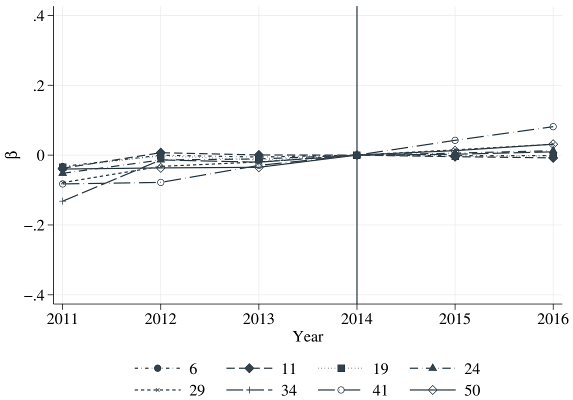
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.49: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. Full controls.

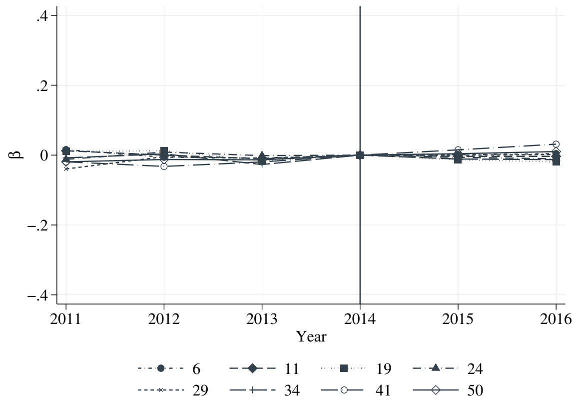


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

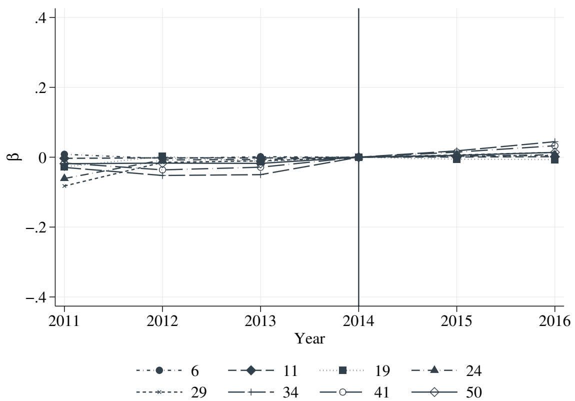
Figure A.50: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: above 16 euros/hour. Full controls.



(a) Bite 1: Regions



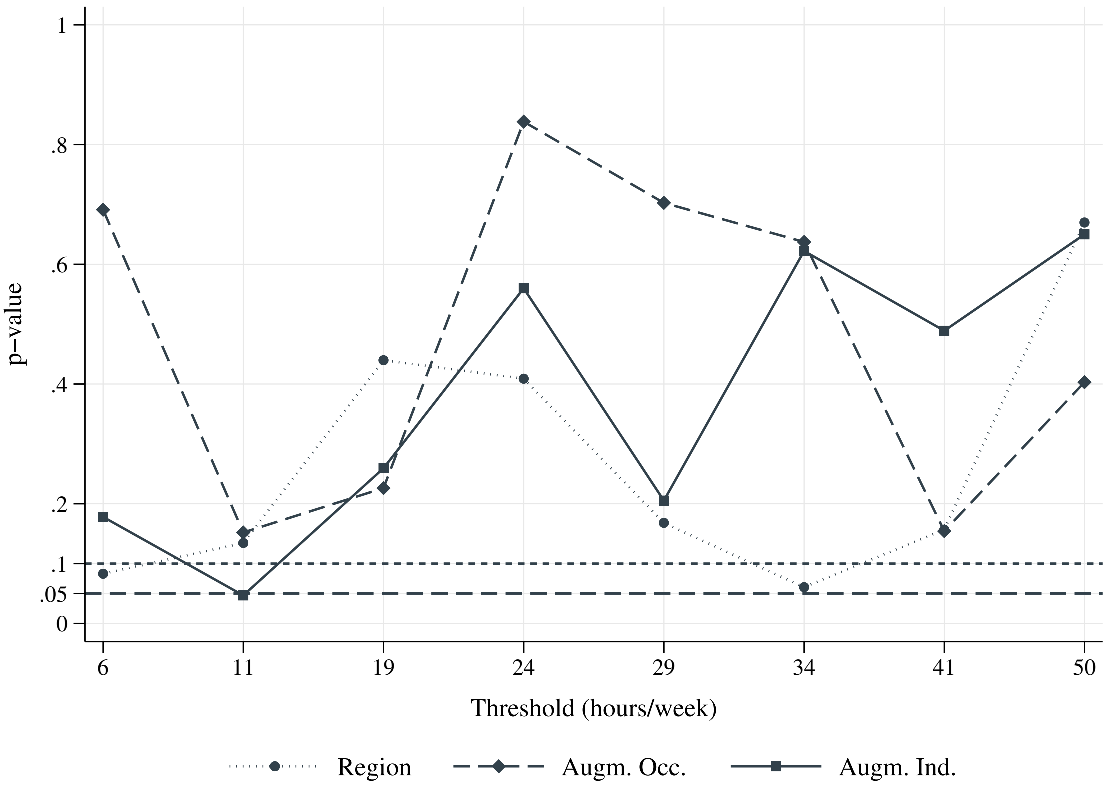
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.51: P-values of joint significance. Weekly hours worked specifications (worker group: above 16 euros/hour). Full controls.

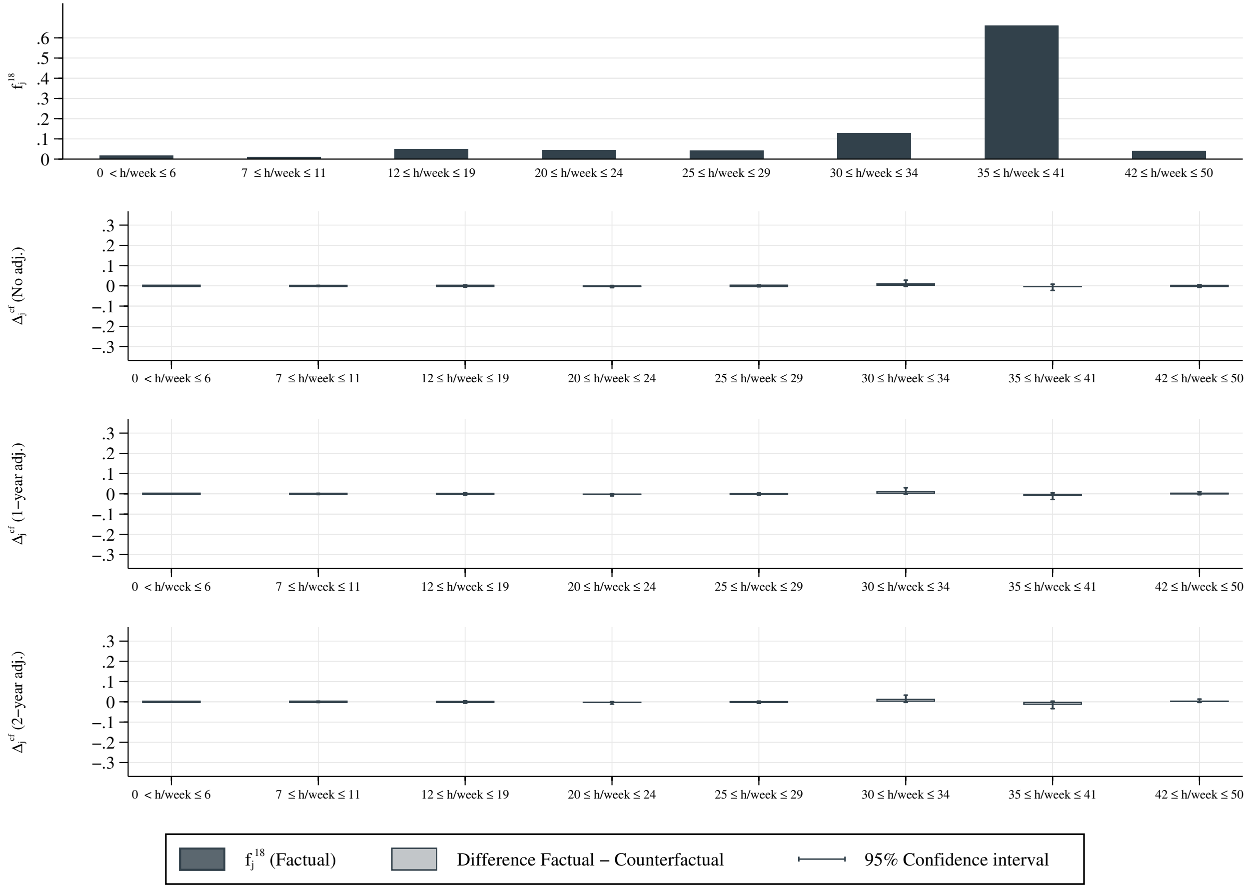


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No firm controls

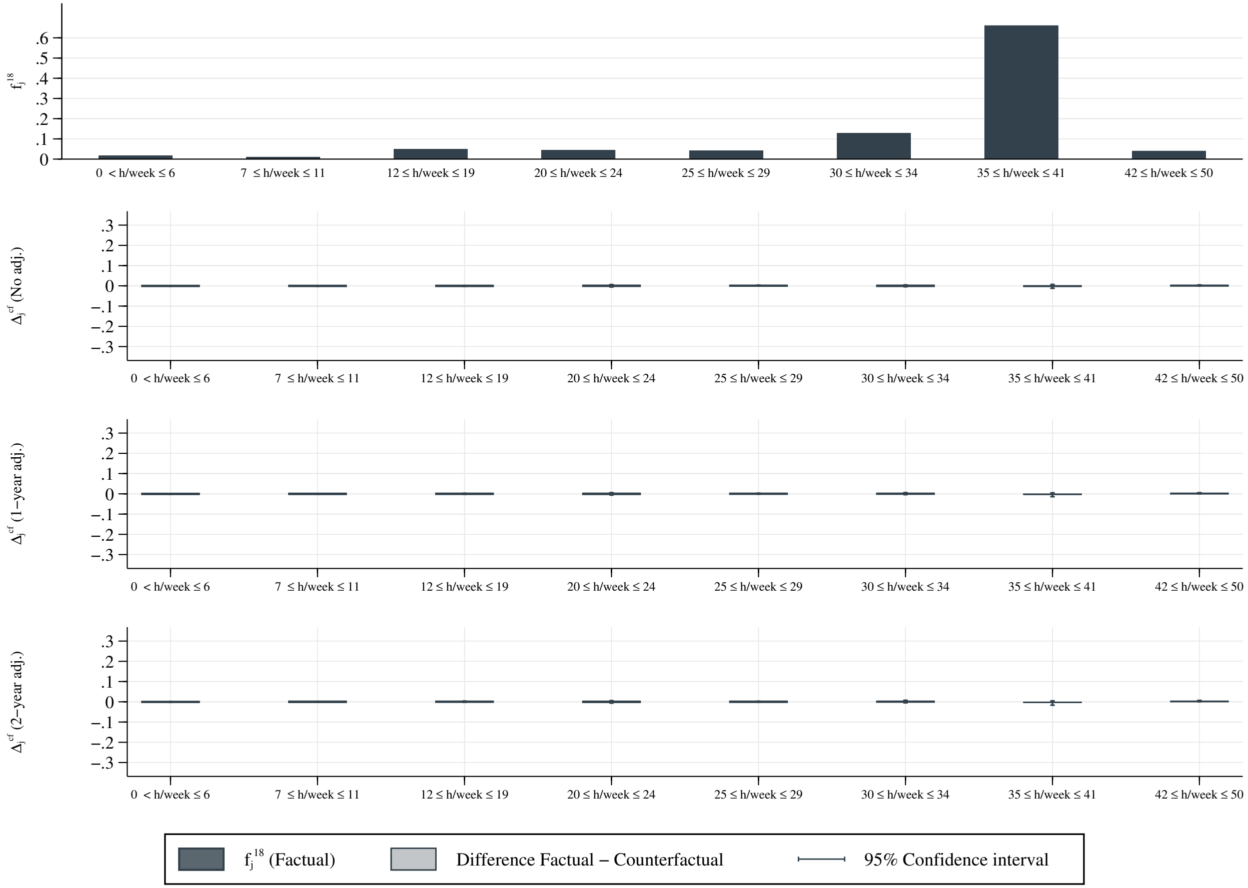
— Intentionally left blank —

Figure A.52: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. No firm controls.



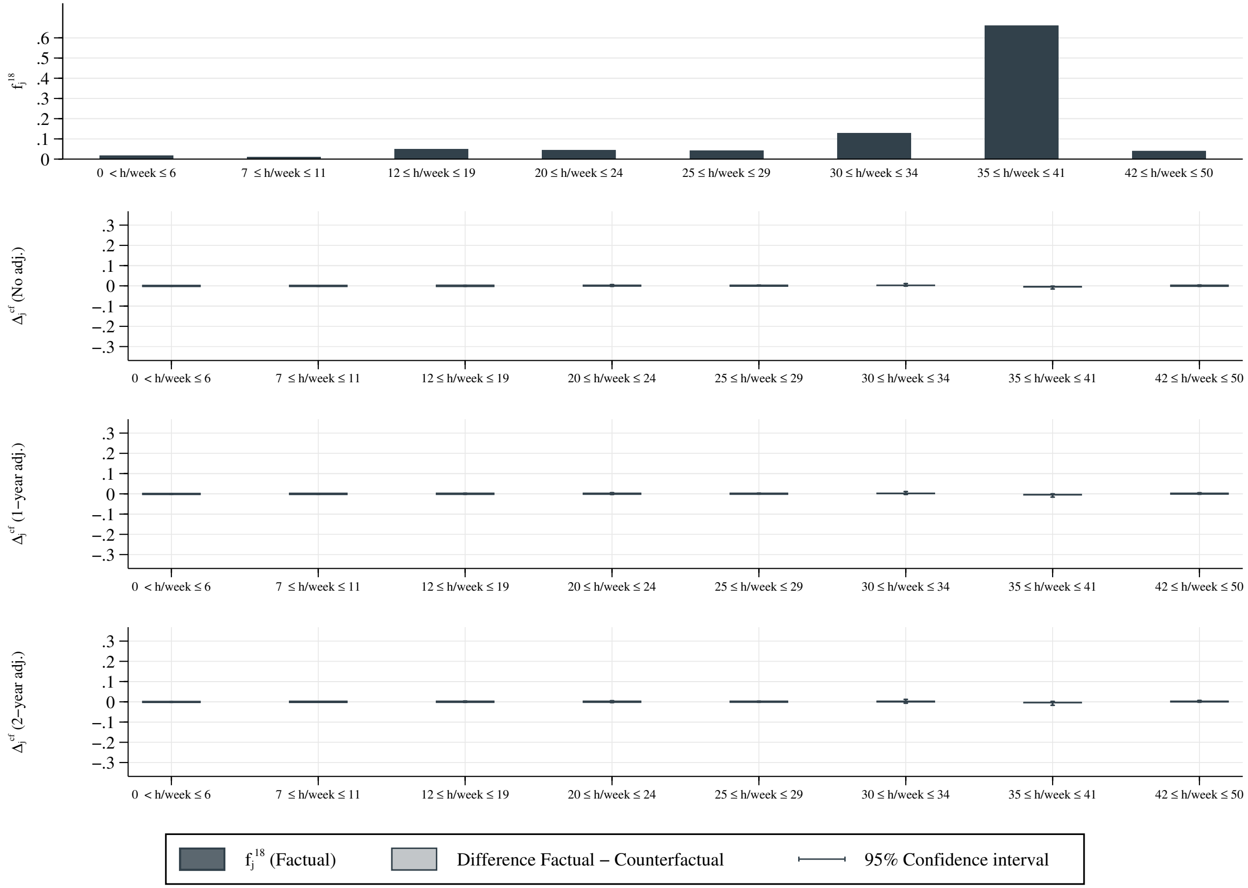
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.53: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. No firm controls.



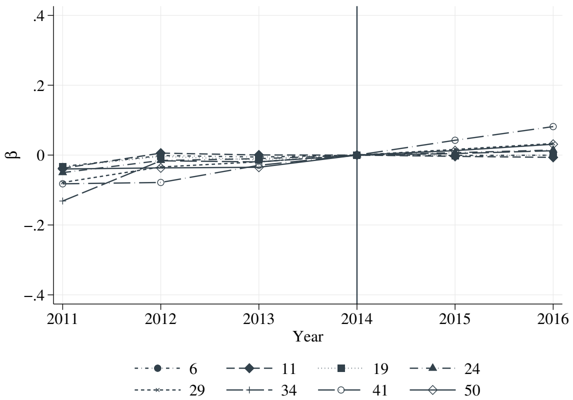
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.54: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. No firm controls.

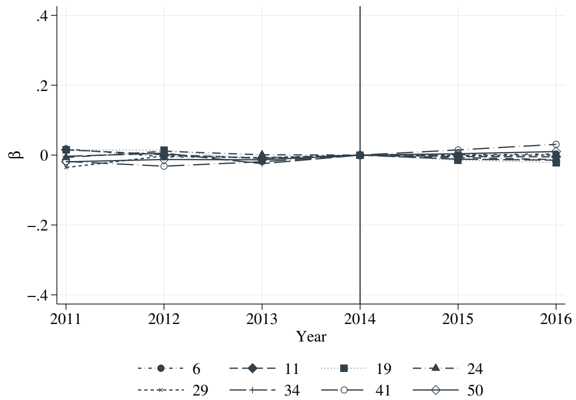


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

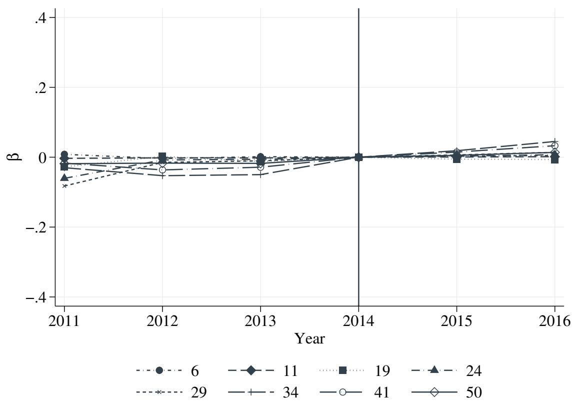
Figure A.55: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: above 16 euros/hour. No firm controls.



(a) Bite 1: Regions



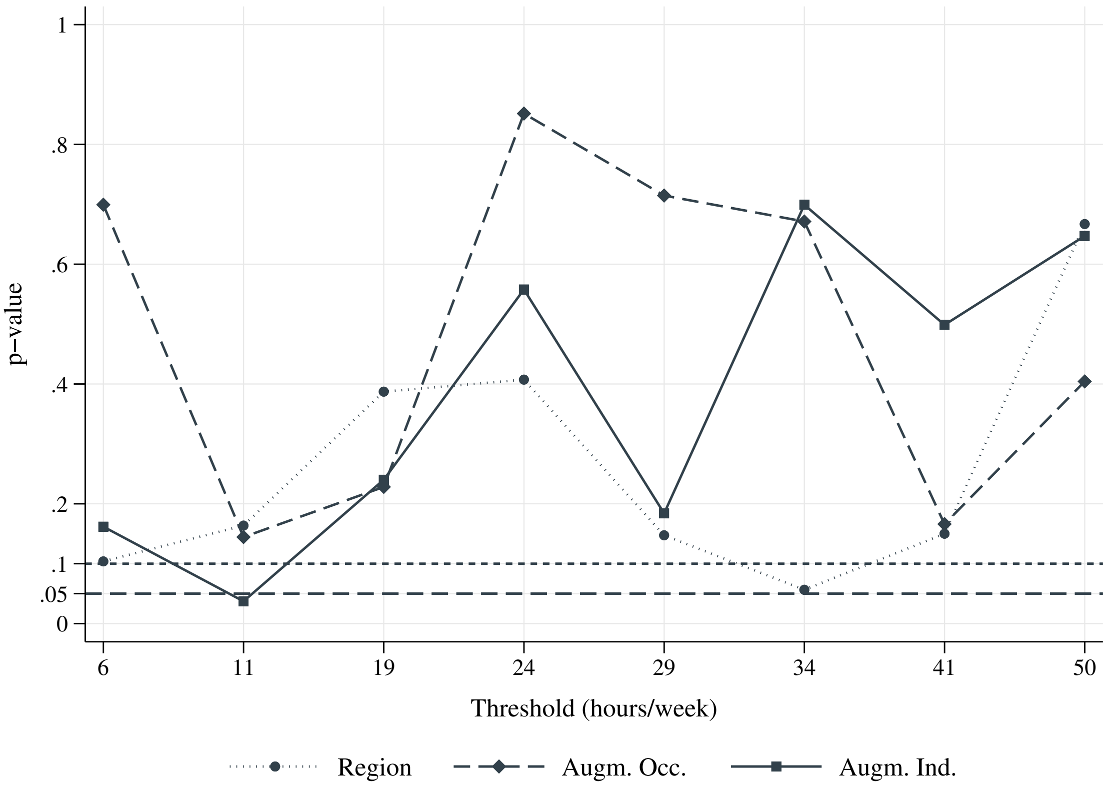
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.56: P-values of joint significance. Weekly hours worked specifications (worker group: above 16 euros/hour). No firm controls.

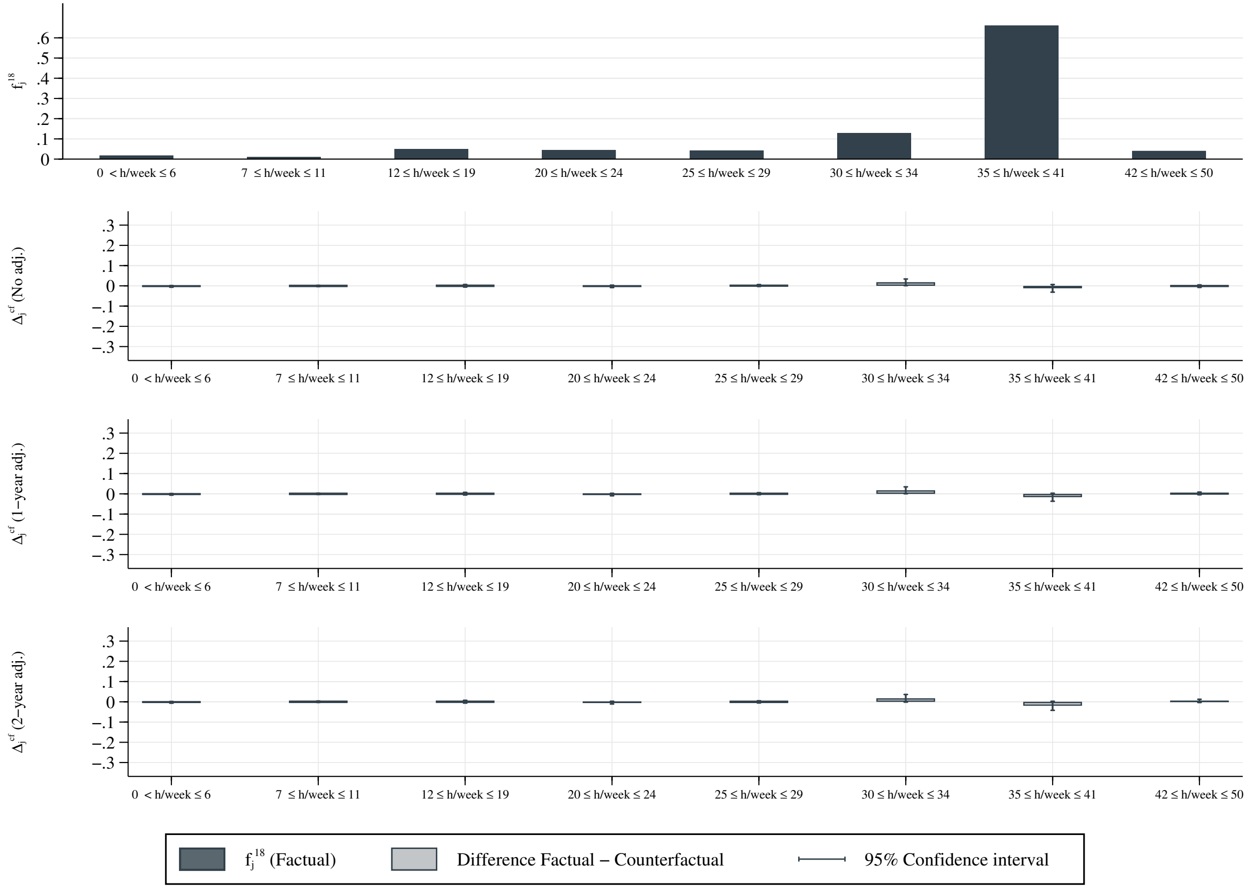


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No controls

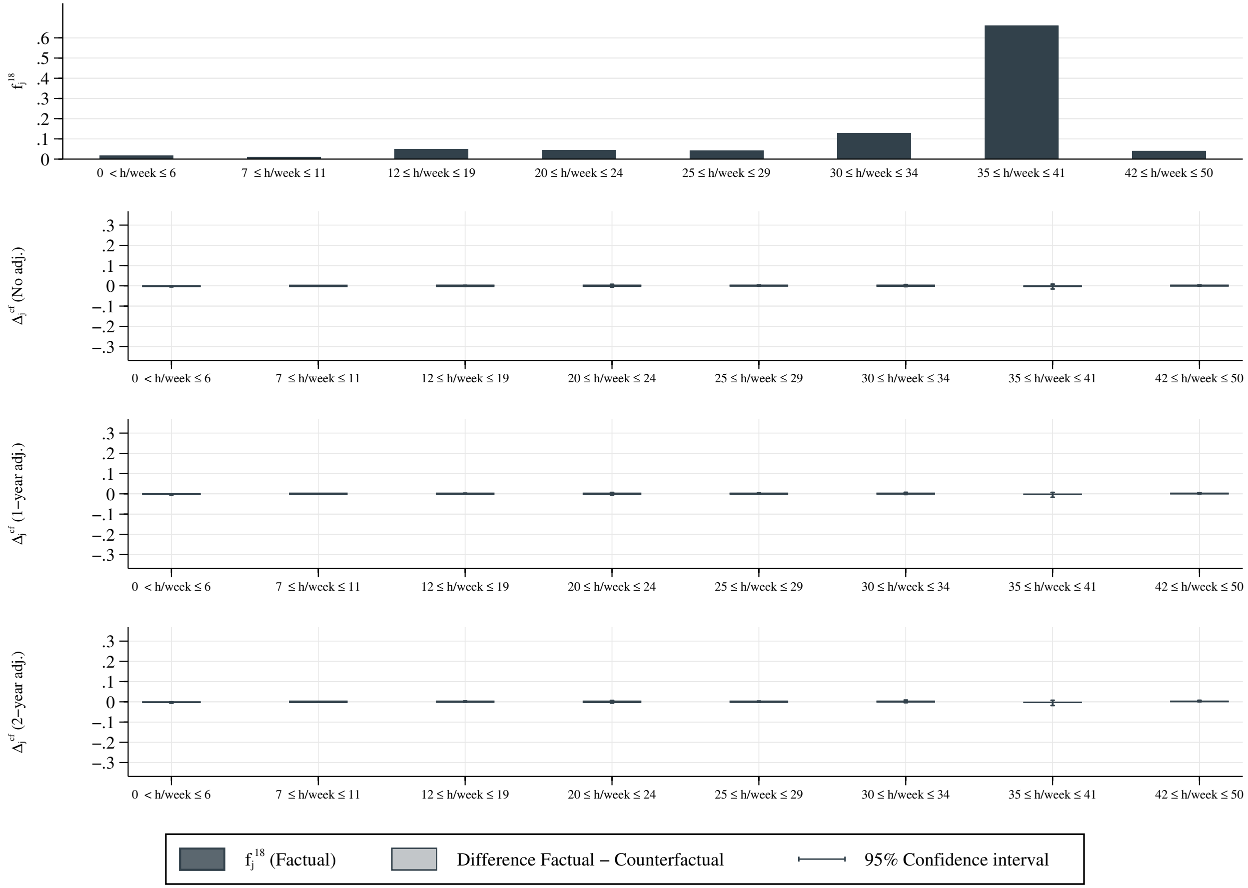
— Intentionally left blank —

Figure A.57: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 1: Regions. All trend adjustment regimes. No controls.



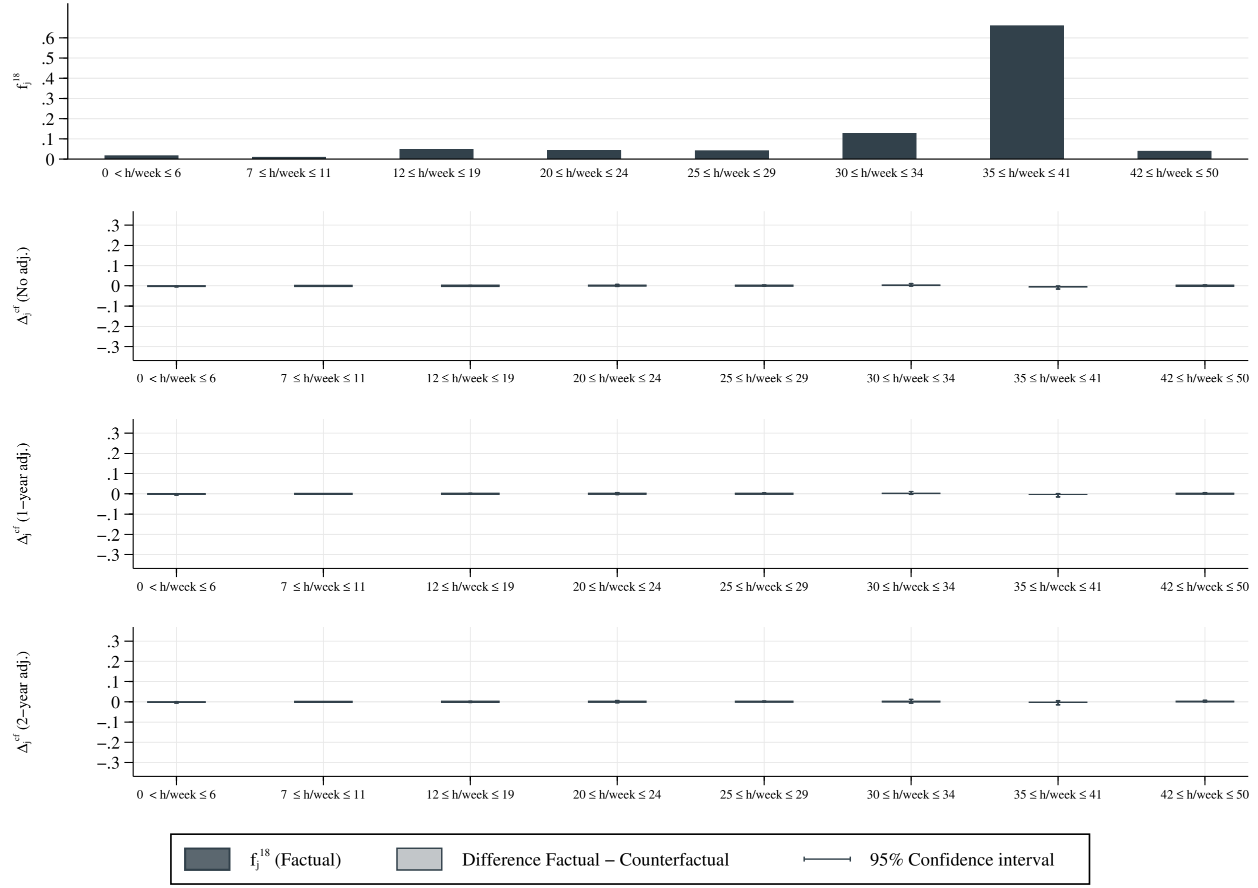
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.58: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 2: Augmented occupations. All trend adjustment regimes. No controls.



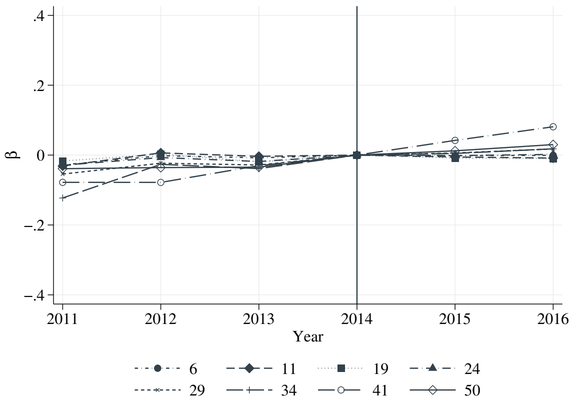
Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.59: 2018 Factual distribution of weekly working hours and treatment effects due to minimum wage for individuals with hourly wages above 16 euros/hour. Bite 3: Augmented industries. All trend adjustment regimes. No controls.

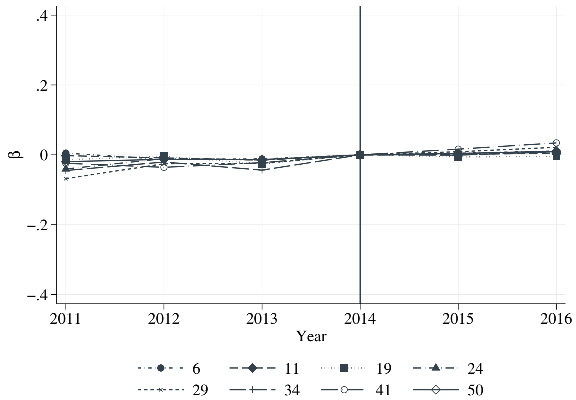


Notes: The bars in the first panel show the factual distributional mass in 2018. The three lower panels show different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

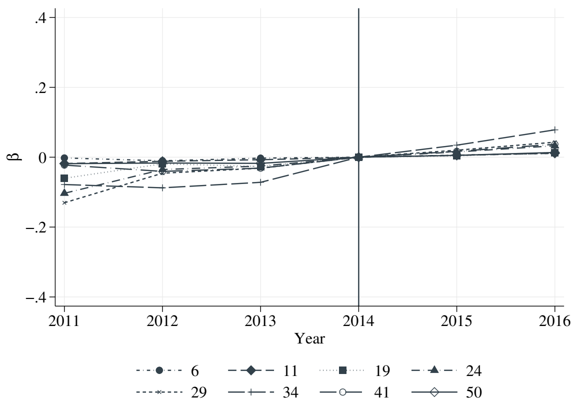
Figure A.60: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
Weekly working hours, worker group: above 16 euros/hour. No controls.



(a) Bite 1: Regions



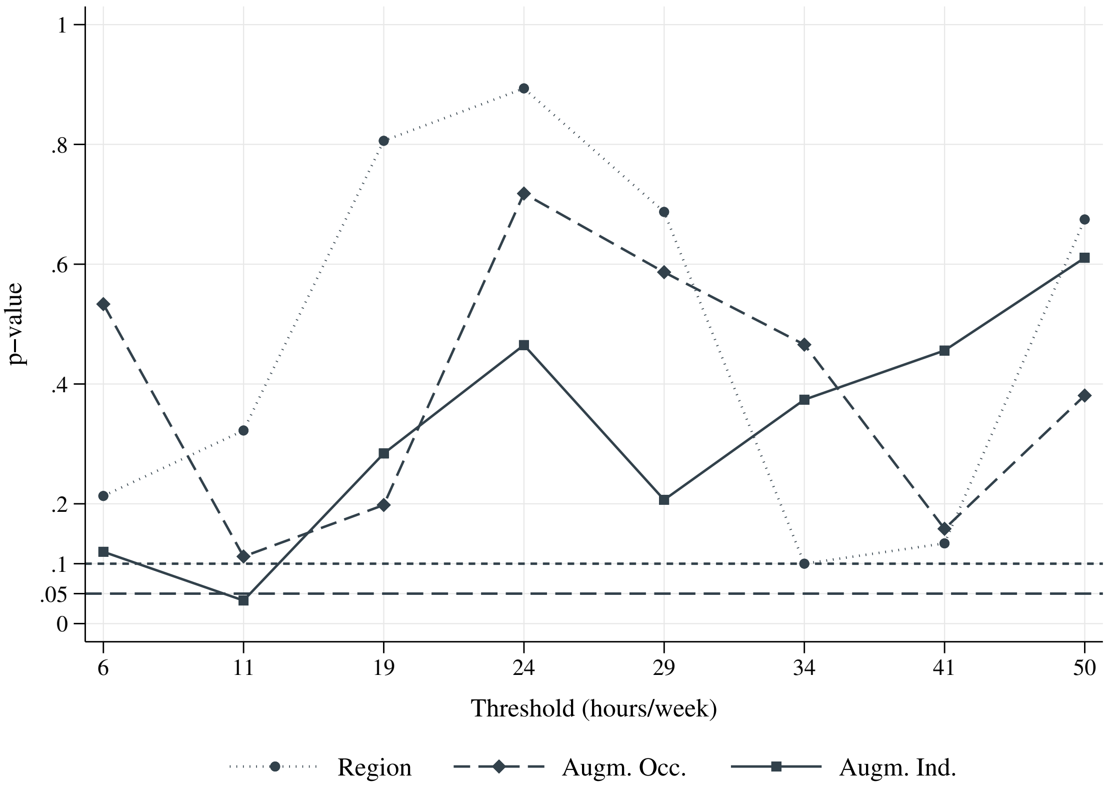
(b) Bite 2: Augmented occupations



(c) Bite 3: Augmented industries

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11) for all weekly hours bin considered in the analyses. Subfigures refer to different bites that have been used in the analyses. Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.61: P-values of joint significance. Weekly hours worked specifications (worker group: above 16 euros/hour). No controls.



Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

A.8.3 Monthly earnings

Table A.8: Minimum wage effects on inequality in monthly earnings, 2014 vs. 2018 – Full controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	2305.181 (32.207)	0.355 (0.002)	11.566 (0.342)	2.131 (0.032)	5.427 (0.117)
2018	2483.971 (27.681)	0.336 (0.002)	10.991 (0.302)	2.108 (0.031)	5.215 (0.101)
$\hat{\Delta}^{18-14}$	178.791*** (43.967)	−0.020*** (0.002)	−0.575 (0.465)	−0.024 (0.046)	−0.212 (0.157)
Bite 1 (Regions)					
No trend adjustment	56.948*** (9.801)	−0.020*** (0.002)	−0.519*** (0.137)	−0.194*** (0.035)	0.215*** (0.072)
1-year trend adjustment	48.948*** (10.268)	−0.016*** (0.002)	−0.615*** (0.165)	−0.179*** (0.032)	0.139* (0.078)
2-year trend adjustment	39.683*** (10.088)	−0.012*** (0.002)	−0.676*** (0.211)	−0.165*** (0.030)	0.081 (0.095)
Bite 2 (Augmented occupations)					
No trend adjustment	−2.521 (14.507)	−0.012*** (0.003)	−0.906*** (0.334)	−0.076 (0.050)	−0.233* (0.130)
1-year trend adjustment	−0.410 (14.976)	−0.009*** (0.003)	−0.644** (0.256)	−0.087* (0.050)	−0.088 (0.102)
2-year trend adjustment	−1.768 (19.378)	−0.006** (0.003)	−0.409 (0.296)	−0.096* (0.050)	0.041 (0.127)
Bite 3 (Augmented industries)					
No trend adjustment	2.222 (8.641)	−0.012*** (0.002)	−0.917*** (0.233)	−0.075* (0.043)	−0.242*** (0.083)
1-year trend adjustment	2.835 (8.602)	−0.009*** (0.002)	−0.695*** (0.205)	−0.083** (0.041)	−0.121* (0.072)
2-year trend adjustment	2.920 (12.785)	−0.006*** (0.002)	−0.492** (0.205)	−0.089** (0.040)	−0.012 (0.078)

Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Table A.9: Minimum wage effects on inequality in monthly earnings, 2014 vs. 2018 – No firm controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	2305.181	0.355	11.566	2.131	5.427
	(32.207)	(0.002)	(0.342)	(0.032)	(0.117)
2018	2483.971	0.336	10.991	2.108	5.215
	(27.681)	(0.002)	(0.302)	(0.031)	(0.101)
$\hat{\Delta}^{18-14}$	178.791***	−0.020***	−0.575	−0.024	−0.212
	(43.967)	(0.002)	(0.465)	(0.046)	(0.157)
Bite 1 (Regions)					
No trend adjustment	60.351***	−0.021***	−0.507***	−0.199***	0.229***
	(10.403)	(0.002)	(0.125)	(0.034)	(0.071)
1-year trend adjustment	53.031***	−0.017***	−0.614***	−0.184***	0.151**
	(11.212)	(0.002)	(0.155)	(0.030)	(0.076)
2-year trend adjustment	44.305***	−0.013***	−0.687***	−0.170***	0.088
	(11.111)	(0.002)	(0.205)	(0.029)	(0.094)
Bite 2 (Augmented occupations)					
No trend adjustment	−0.238	−0.013***	−0.910***	−0.080	−0.225*
	(14.782)	(0.003)	(0.336)	(0.050)	(0.133)
1-year trend adjustment	0.716	−0.009***	−0.633**	−0.089*	−0.077
	(15.243)	(0.003)	(0.257)	(0.049)	(0.103)
2-year trend adjustment	−0.351	−0.006**	−0.388	−0.097**	0.053
	(19.259)	(0.003)	(0.306)	(0.049)	(0.130)
Bite 3 (Augmented industries)					
No trend adjustment	2.223	−0.012***	−0.921***	−0.075*	−0.242***
	(8.903)	(0.002)	(0.231)	(0.042)	(0.085)
1-year trend adjustment	2.669	−0.009***	−0.695***	−0.083**	−0.119
	(8.806)	(0.002)	(0.203)	(0.040)	(0.074)
2-year trend adjustment	2.723	−0.006***	−0.491**	−0.090**	−0.010
	(13.137)	(0.002)	(0.204)	(0.039)	(0.079)

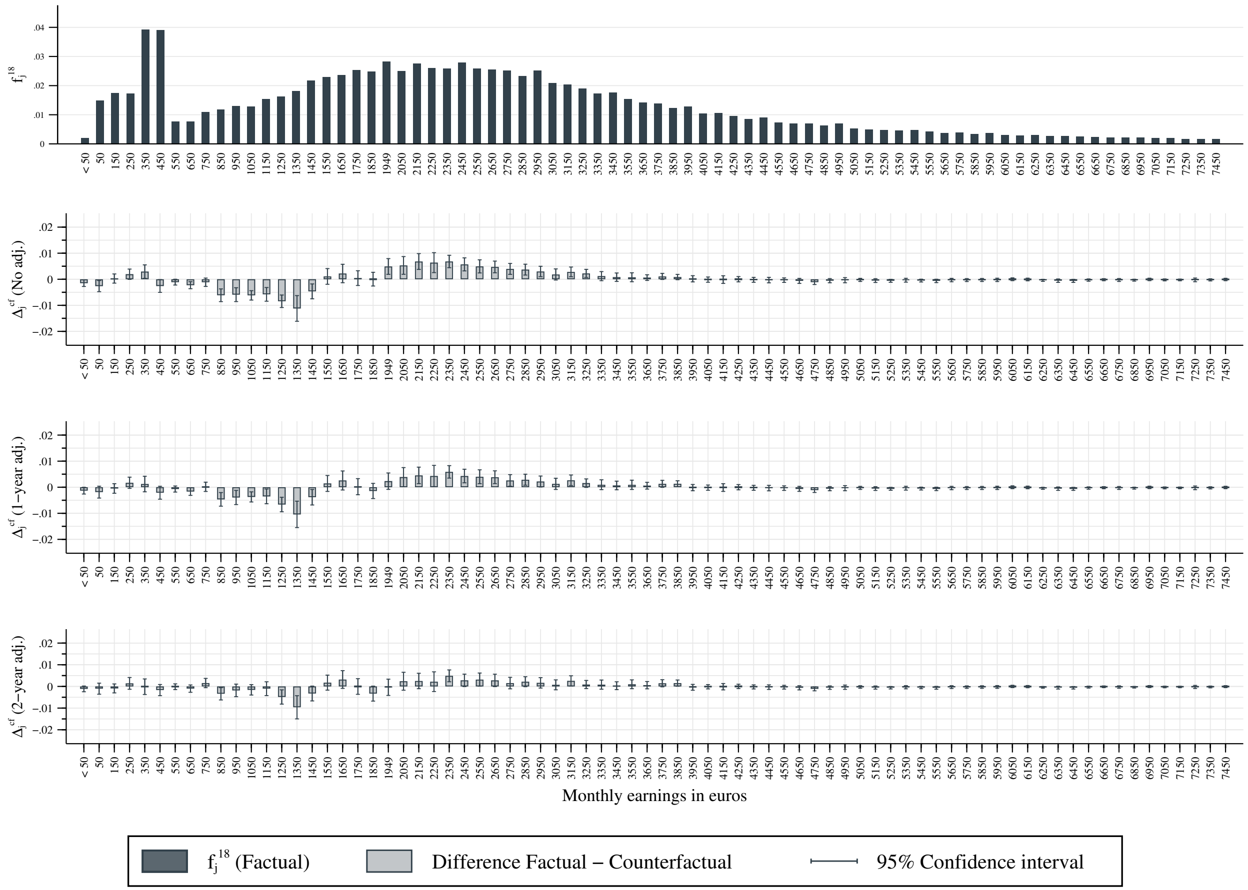
Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Table A.10: Minimum wage effects on inequality in monthly earnings, 2014 vs. 2018s – No additional controls

	Mean	Gini	Q90/Q10	Q90/Q50	Q50/Q10
2014	2305.181	0.355	11.566	2.131	5.427
	(32.207)	(0.002)	(0.342)	(0.032)	(0.117)
2018	2483.971	0.336	10.991	2.108	5.215
	(27.681)	(0.002)	(0.302)	(0.031)	(0.101)
$\hat{\Delta}^{18-14}$	178.791***	−0.020***	−0.575	−0.024	−0.212
	(43.967)	(0.002)	(0.465)	(0.046)	(0.157)
Bite 1 (Regions)					
No trend adjustment	84.794***	−0.025***	−0.896***	−0.226***	0.122
	(19.067)	(0.004)	(0.312)	(0.038)	(0.125)
1-year trend adjustment	80.052***	−0.022***	−1.092***	−0.211***	0.004
	(19.531)	(0.004)	(0.331)	(0.036)	(0.131)
2-year trend adjustment	73.245***	−0.019***	−1.239***	−0.196***	−0.094
	(20.013)	(0.004)	(0.364)	(0.034)	(0.144)
Bite 2 (Augmented occupations)					
No trend adjustment	15.168	−0.018***	−1.467***	−0.096*	−0.439*
	(15.821)	(0.005)	(0.523)	(0.056)	(0.237)
1-year trend adjustment	17.895	−0.015***	−1.203**	−0.106*	−0.293
	(16.008)	(0.005)	(0.475)	(0.056)	(0.230)
2-year trend adjustment	18.102	−0.012**	−0.959**	−0.115**	−0.162
	(18.502)	(0.005)	(0.468)	(0.056)	(0.233)
Bite 3 (Augmented industries)					
No trend adjustment	18.286	−0.017***	−1.328***	−0.086*	−0.401**
	(11.944)	(0.003)	(0.361)	(0.045)	(0.173)
1-year trend adjustment	20.052*	−0.014***	−1.119***	−0.094**	−0.286
	(11.713)	(0.003)	(0.360)	(0.043)	(0.174)
2-year trend adjustment	21.459	−0.011***	−0.918**	−0.101**	−0.178
	(14.237)	(0.003)	(0.378)	(0.042)	(0.184)

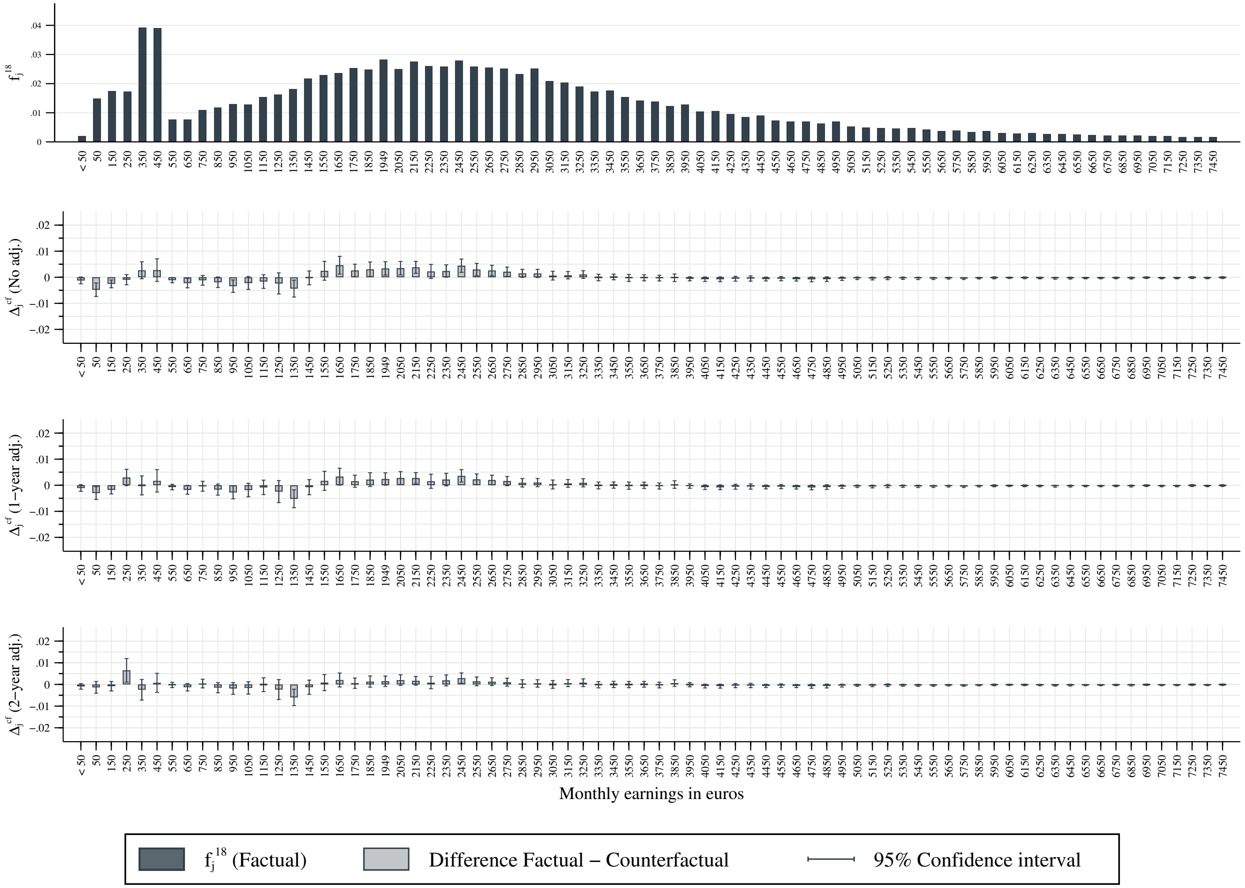
Notes: Estimates in rows four to twelve refer to eq. (2.10). Bootstrap standard errors (100 replications) in parentheses. Bootstrap standard errors for factual values (rows one to three) are clustered at the regional level. Bootstrap standard errors for the counterfactual values and differences are clustered at the respective treatment level (region, augmented occupation or augmented industry level). ***/**/* indicate statistical significance for the factual/counterfactual differences at the 1%/5%/10% level.
Sources: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.62: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 1: Regions. Full controls.



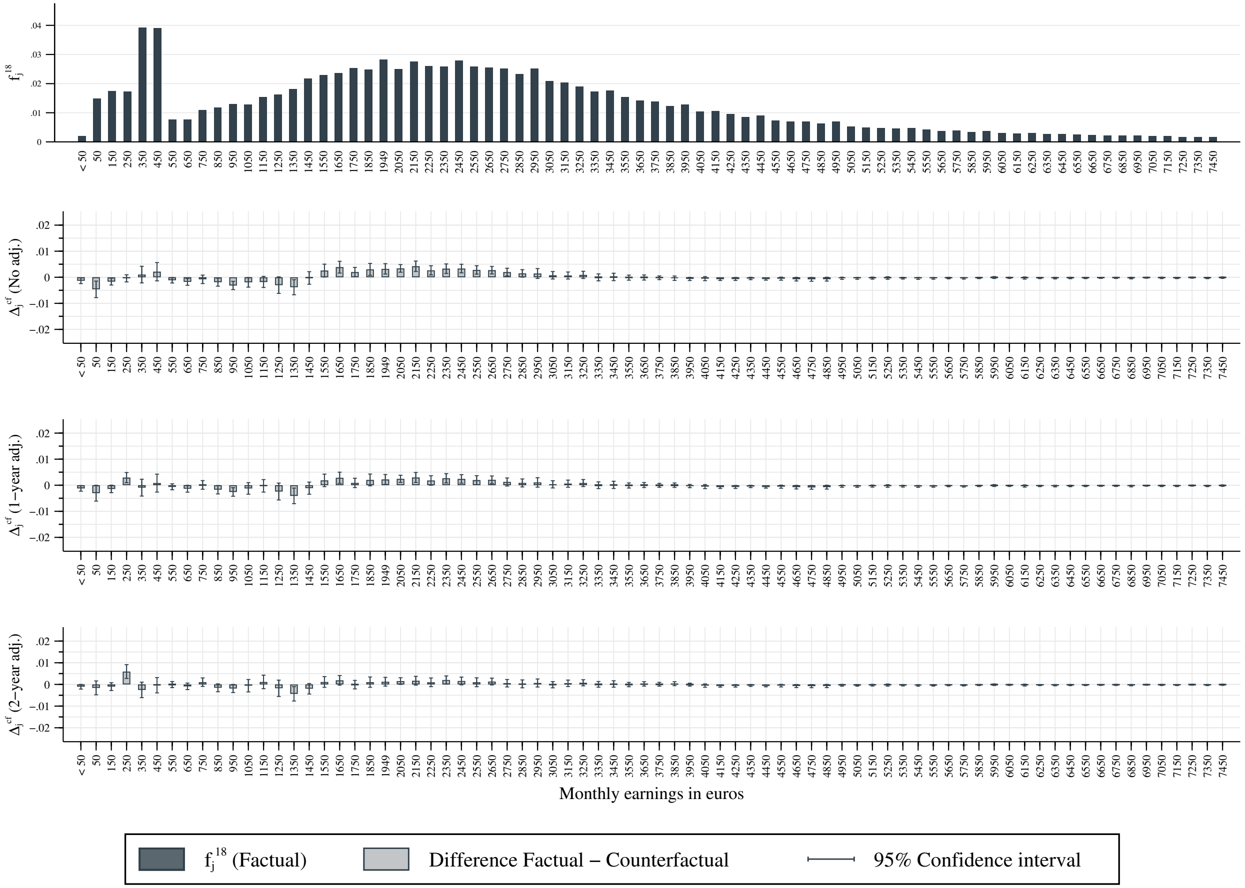
Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.63: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. Full controls.



Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

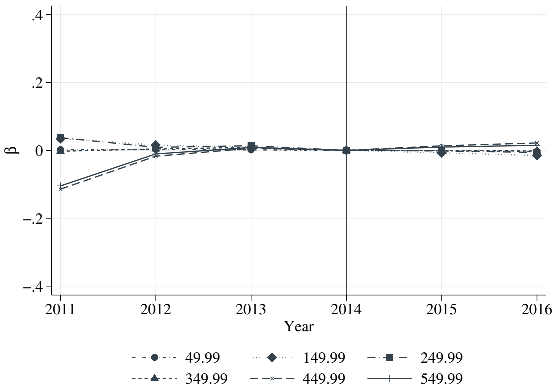
Figure A.64: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. Full controls.



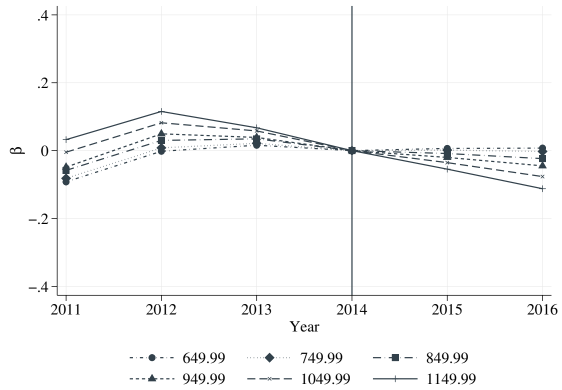
Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

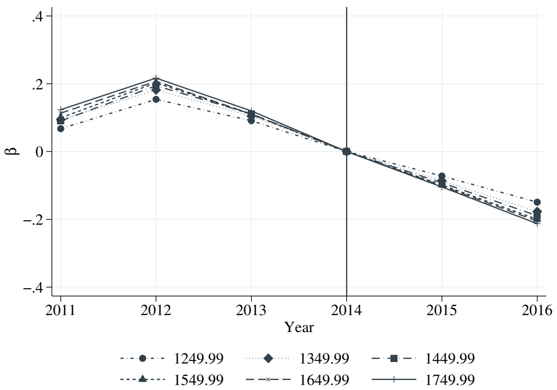
Figure A.65: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (lower to middle thresholds), bite 1 (region). Full controls.



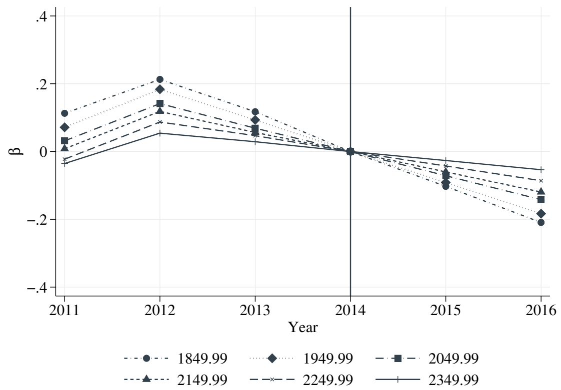
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



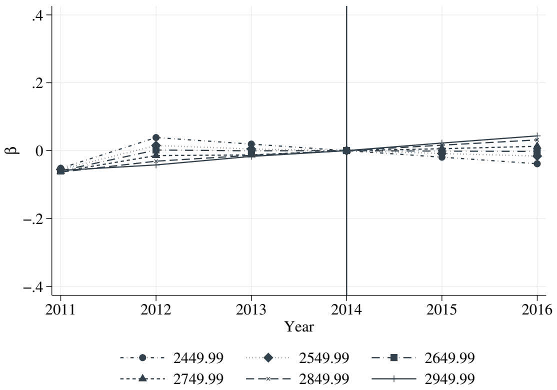
(c) Thresholds: 1,249.99 to 1,749.99



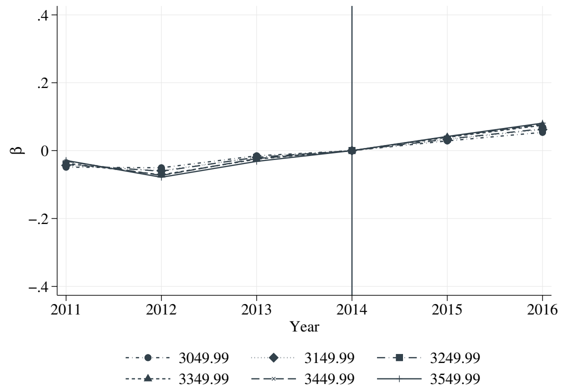
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

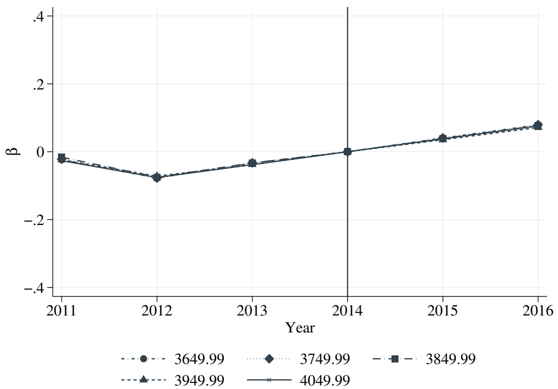
Figure A.66: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 1 (region). Full controls.



(a) Thresholds: 2,449.99 to 2,949.99



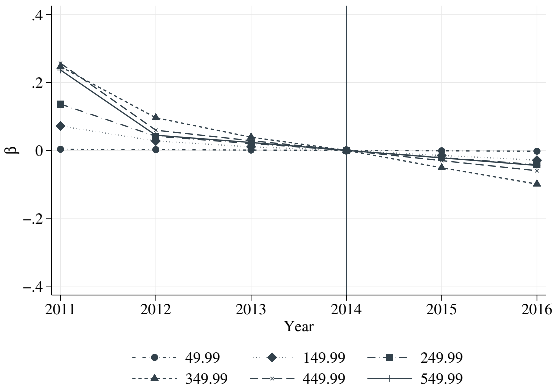
(b) Thresholds: 3,049.99 to 3,549.99



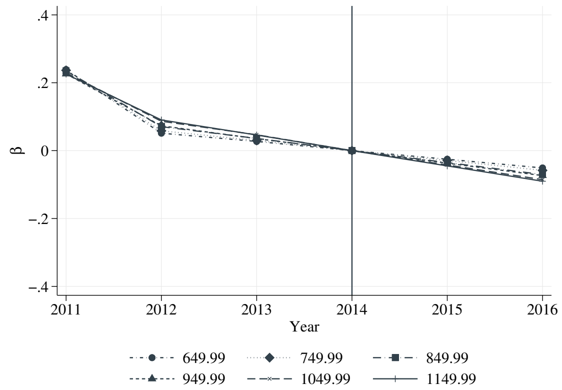
(c) Thresholds: 3,649.99 to 4049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

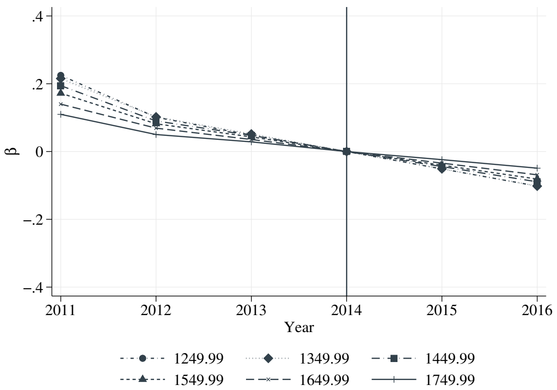
Figure A.67: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
– Monthly earnings (lower to middle thresholds), bite 2 (augmented occupations). Full controls.



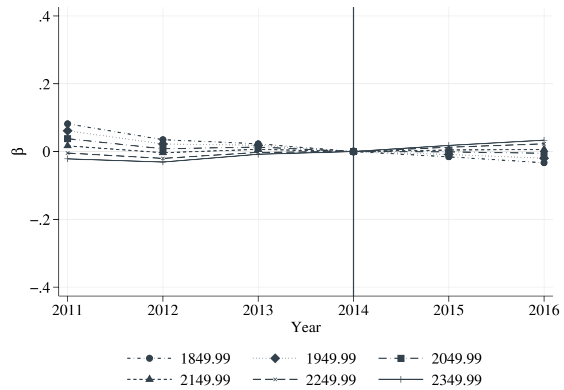
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1149.99



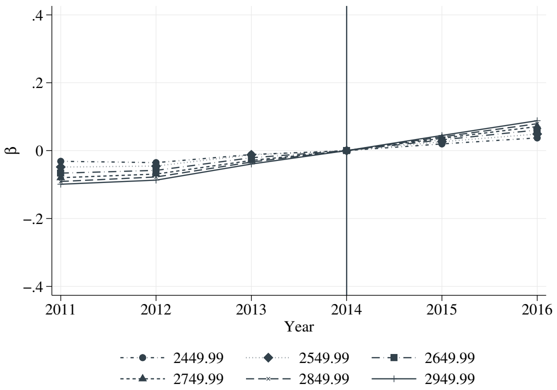
(c) Thresholds: 1,249.99 to 1,749.99



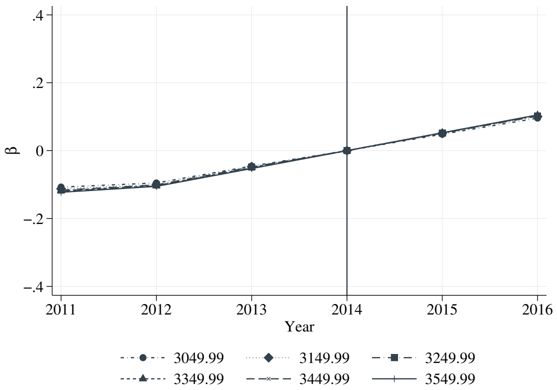
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

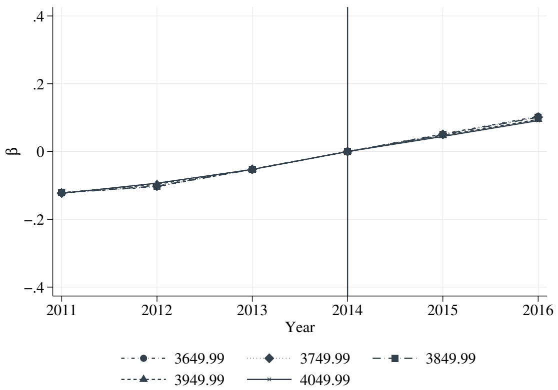
Figure A.68: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 2 (augmented occupations). Full controls.



(a) Thresholds: 2,449.99 to 2,949.99



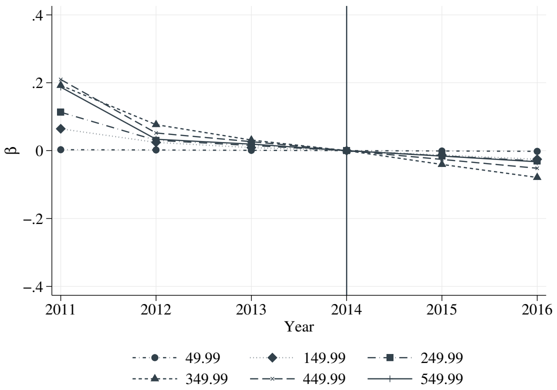
(b) Thresholds: 3,049.99 to 3,549.99



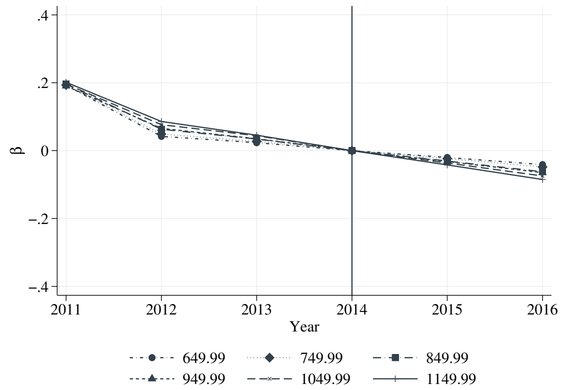
(c) Thresholds: 3,649.99 to 4049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

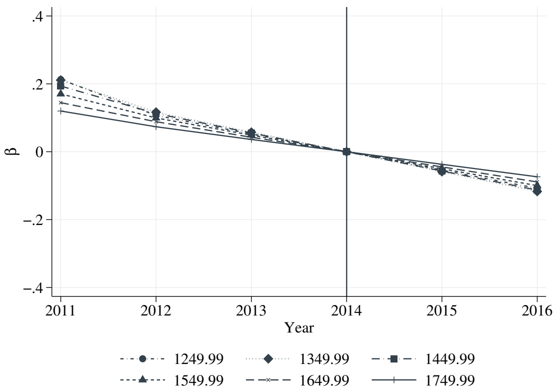
Figure A.69: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
– Monthly earnings (lower to middle thresholds), bite 3 (augmented industries). Full controls.



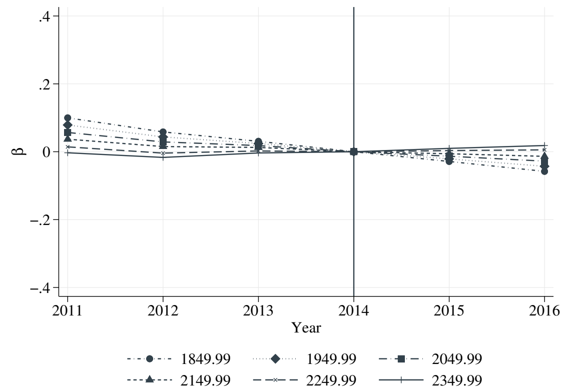
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



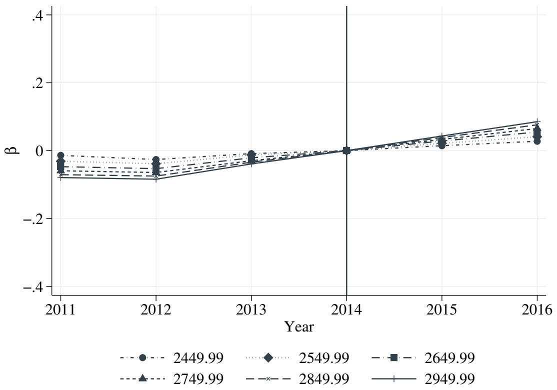
(c) Thresholds: 1,249.99 to 1,749.99



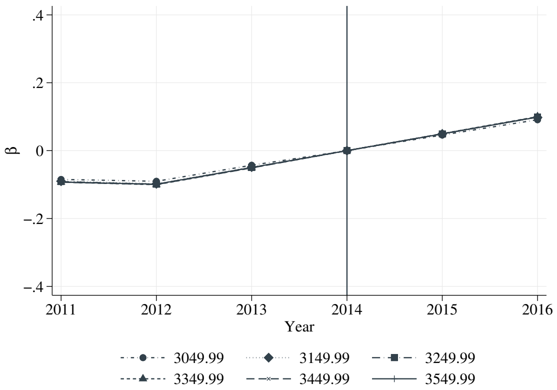
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

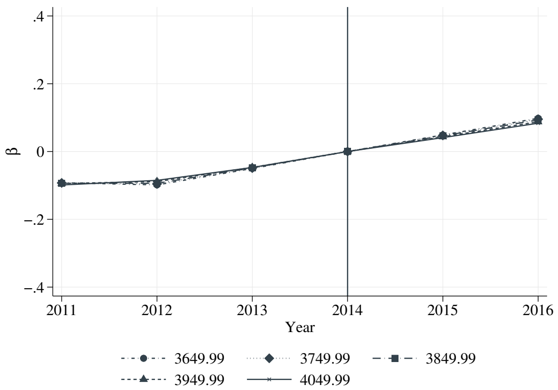
Figure A.70: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 3 (augmented industries). Full controls.



(a) Thresholds: 2,449.99 to 2,949.99



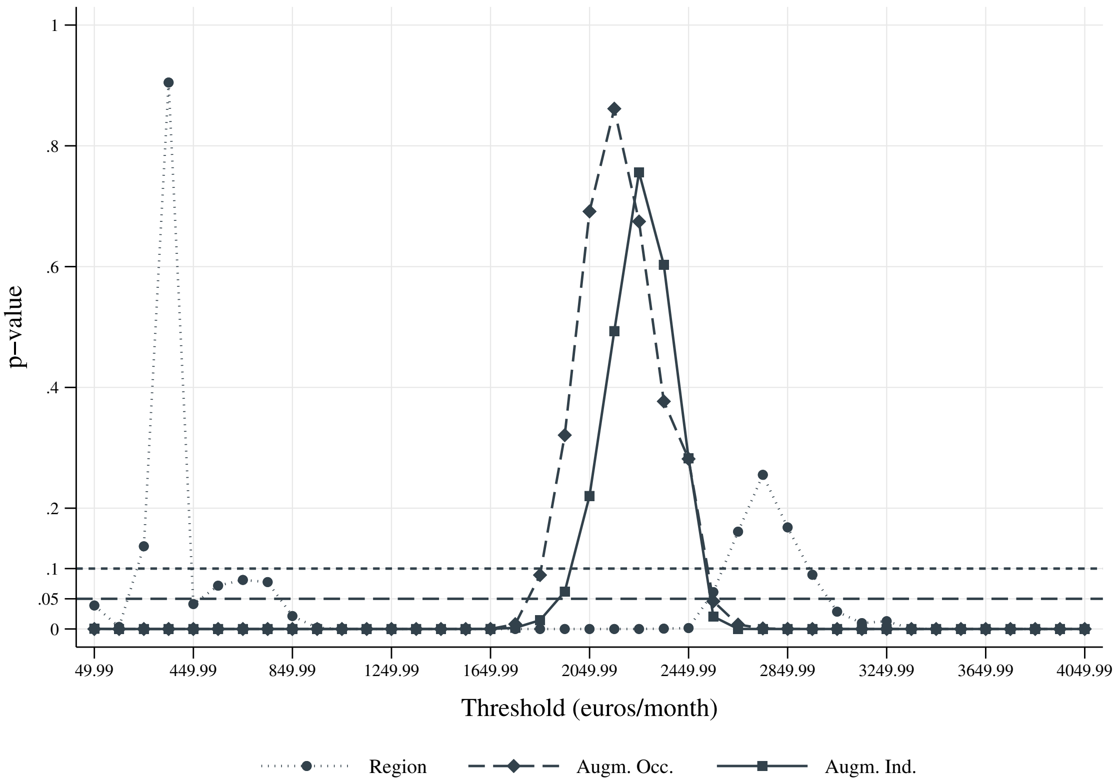
(b) Thresholds: 3,049.99 to 3,549.99



(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.71: P-values of joint significance – Monthly earnings. Full controls.

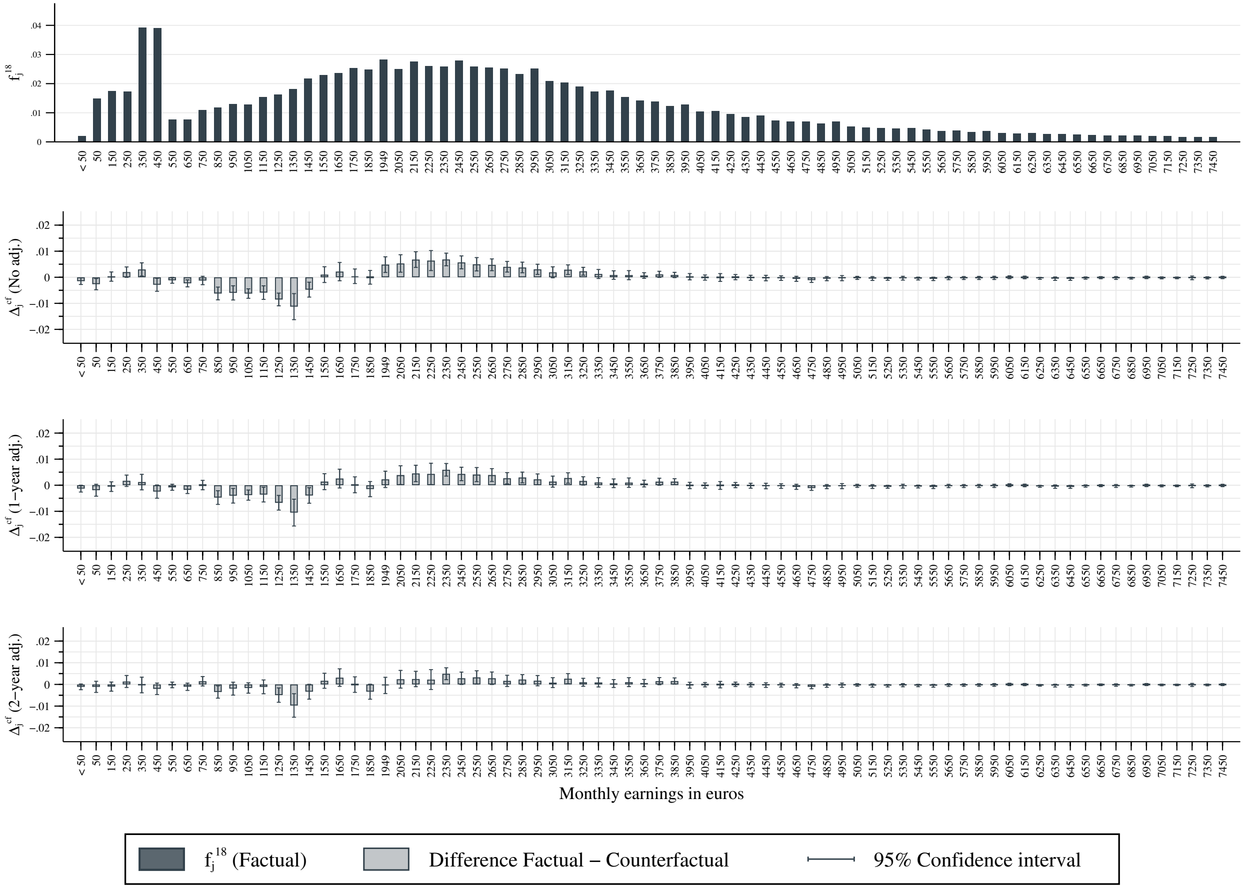


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No firm controls

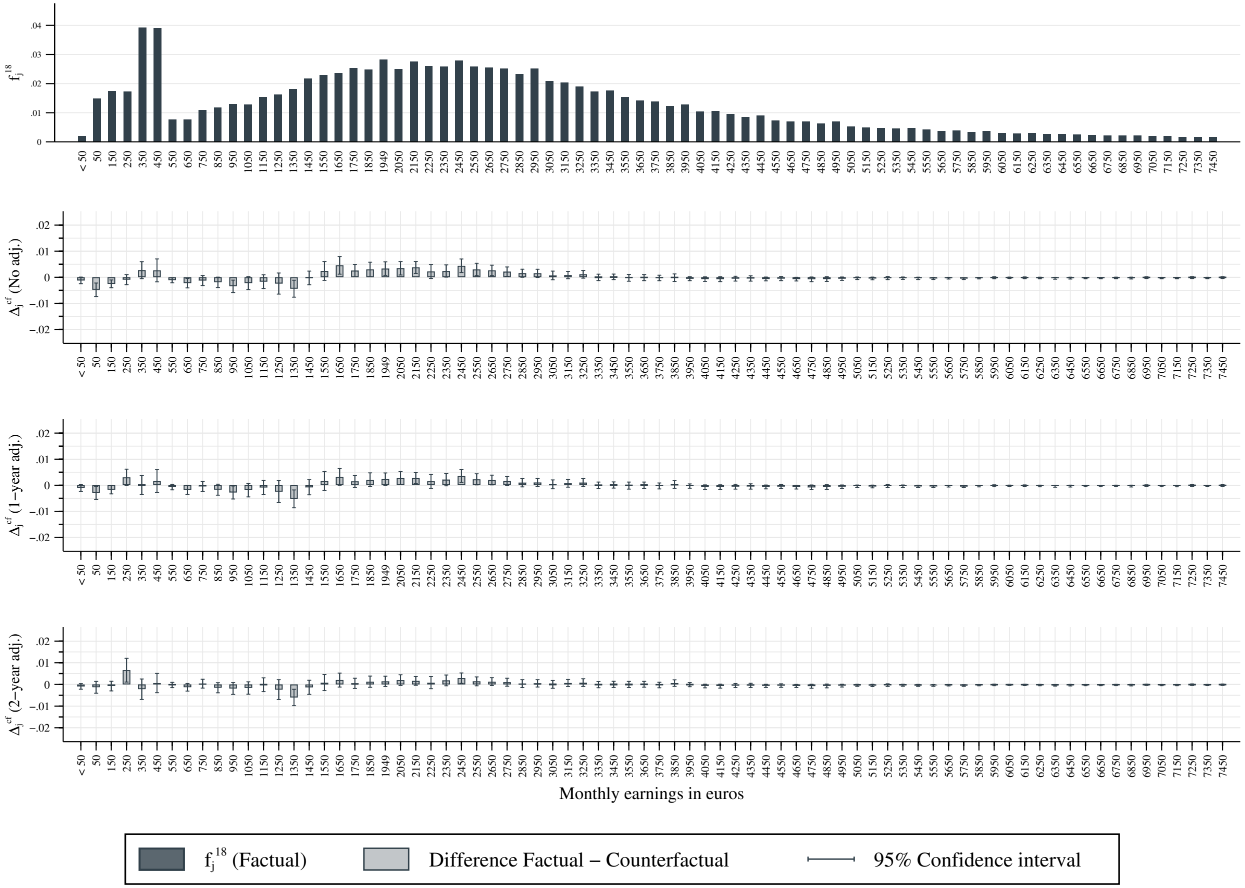
— Intentionally left blank —

Figure A.72: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 1: Regions. No firm controls.



Notes: The x-axis shows monthly wage bins. For example, the ‘1050’ bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

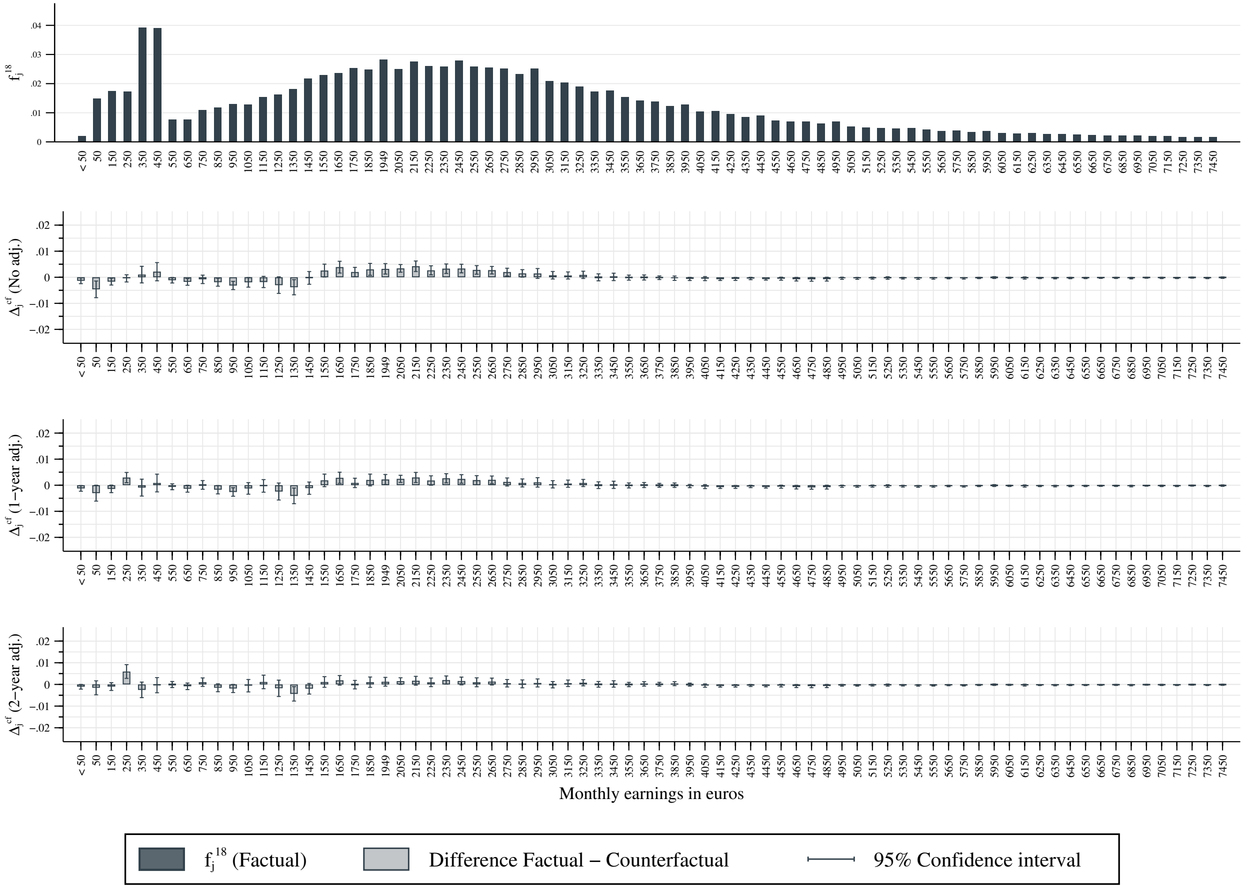
Figure A.73: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. No firm controls.



Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

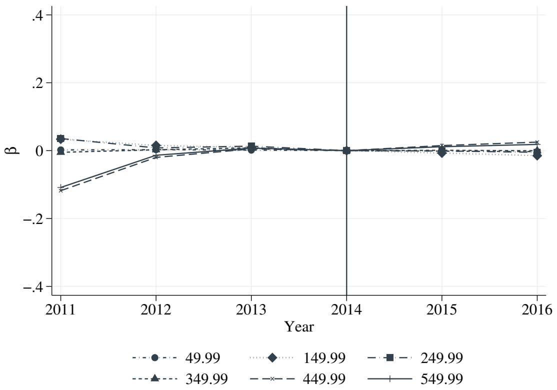
Figure A.74: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. No firm controls.



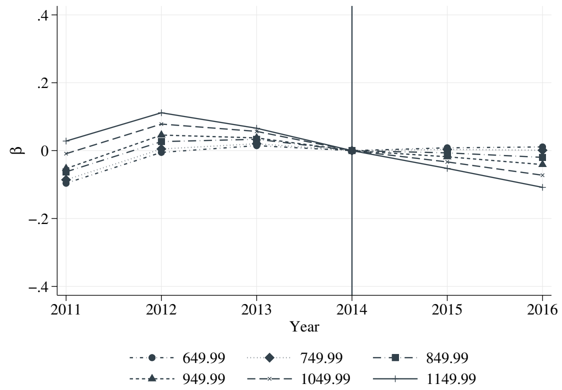
Notes: The x-axis shows monthly wage bins. For example, the ‘1050’ bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).

Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

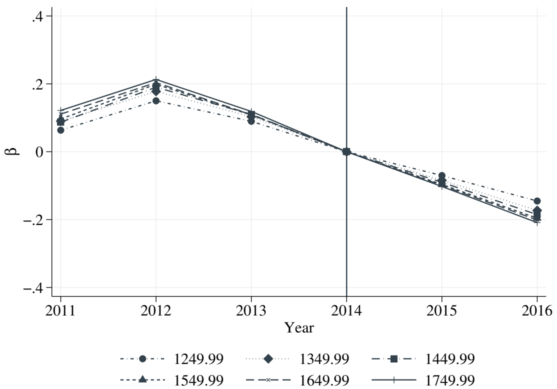
Figure A.75: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (lower to middle thresholds), bite 1 (region). No firm controls.



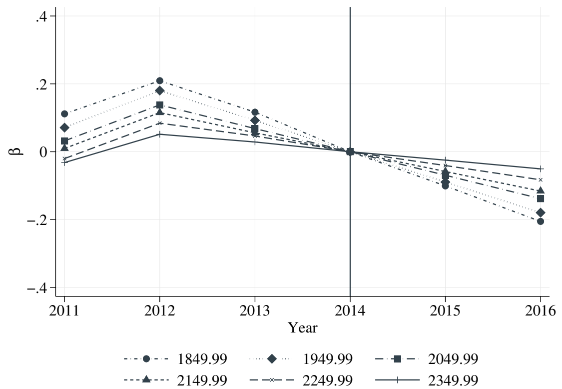
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



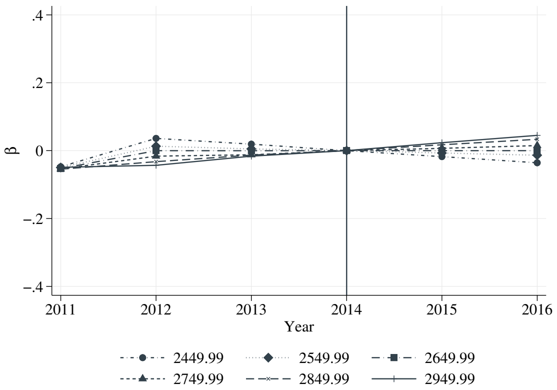
(c) Thresholds: 1,249.99 to 1,749.99



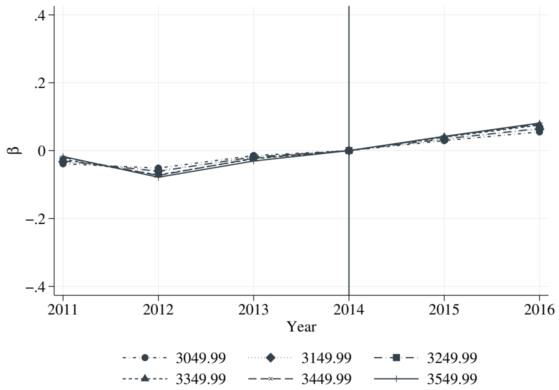
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

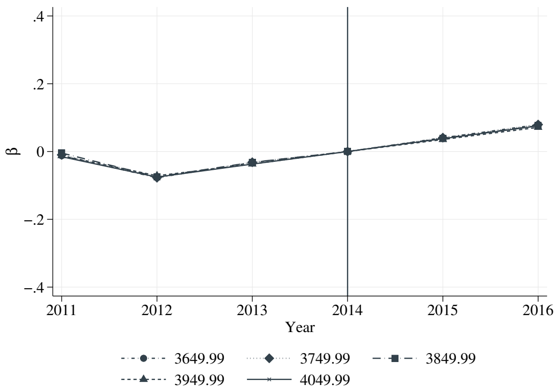
Figure A.76: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 1 (region). No firm controls.



(a) Thresholds: 2,449.99 to 2,949.99



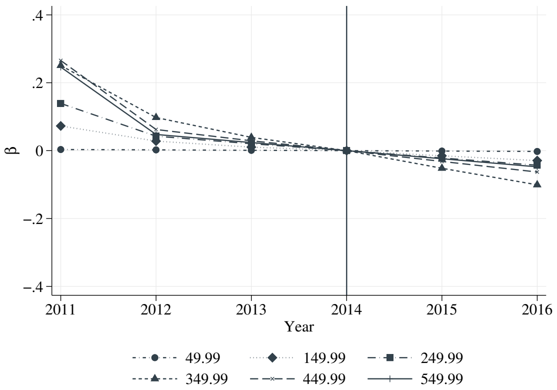
(b) Thresholds: 3,049.99 to 3,549.99



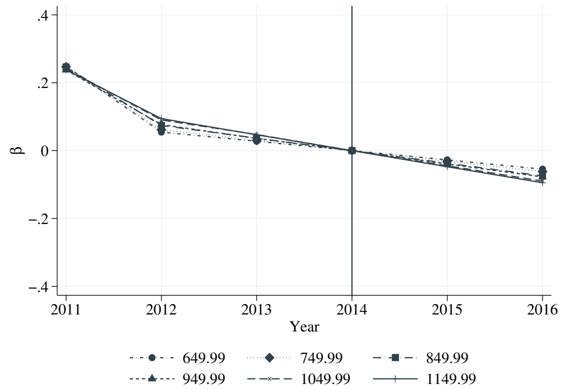
(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

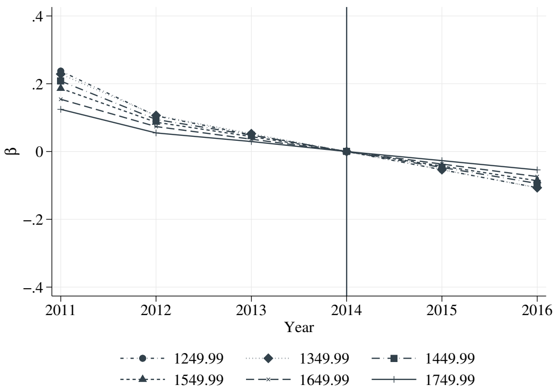
Figure A.77: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (lower to middle thresholds), bite 2 (augmented occupations). No firm controls.



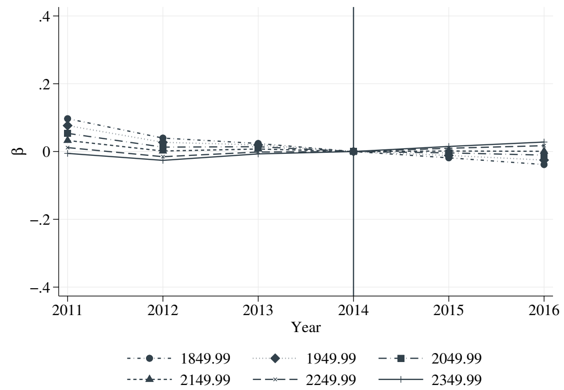
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



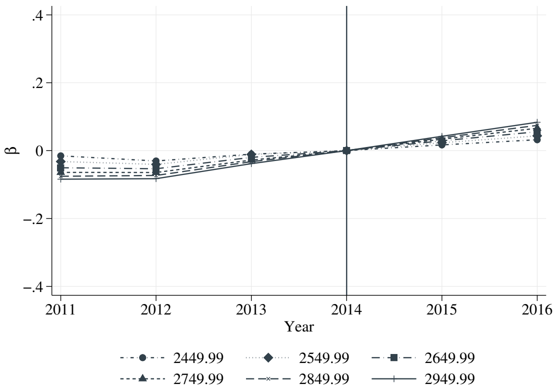
(c) Thresholds: 1,249.99 to 1,749.99



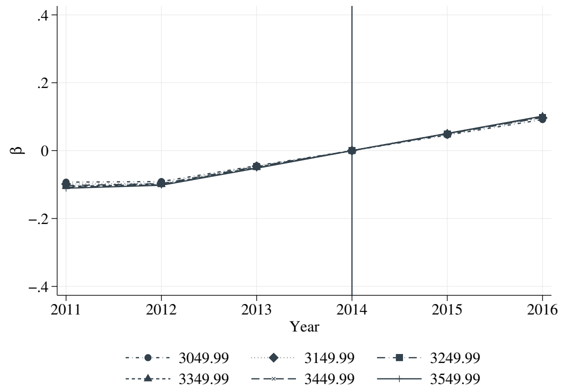
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

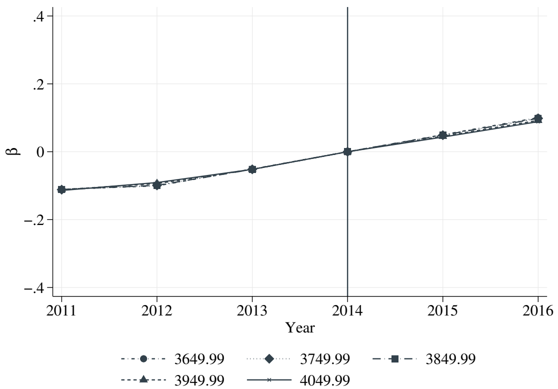
Figure A.78: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 2 (augmented occupations). No firm controls.



(a) Thresholds: 2,449.99 to 2,949.99



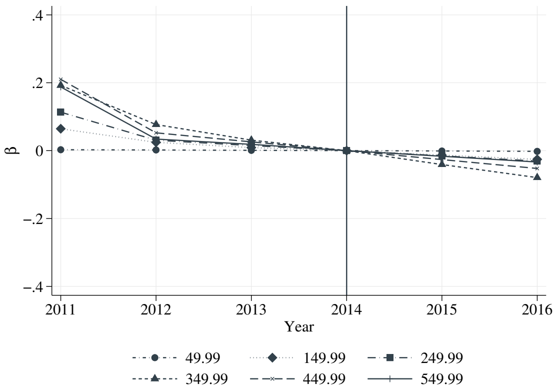
(b) Thresholds: 3,049.99 to 3,549.99



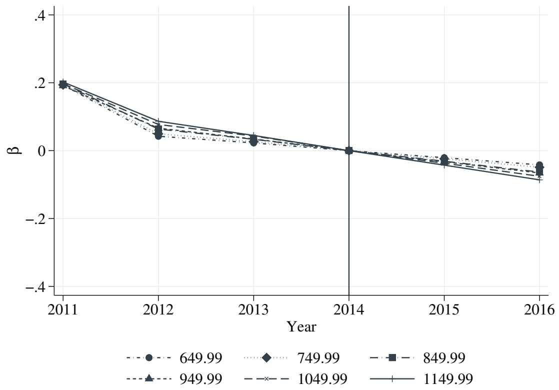
(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

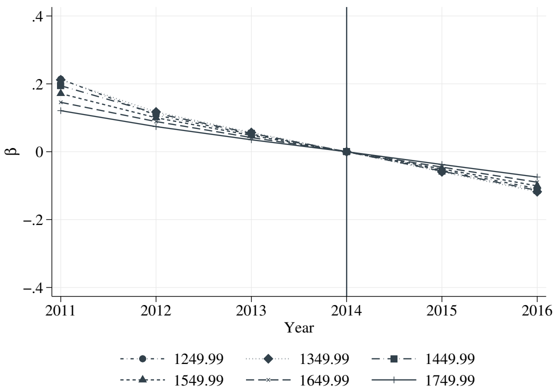
Figure A.79: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
– Monthly earnings (lower to middle thresholds), bite 3 (augmented industries). No firm controls.



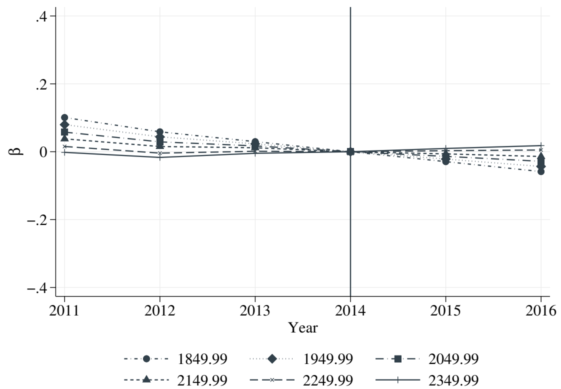
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



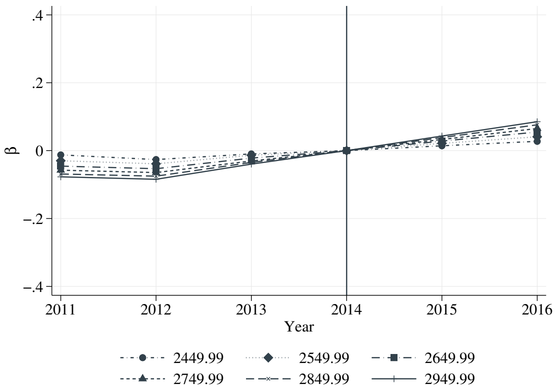
(c) Thresholds: 1,249.99 to 1,749.99



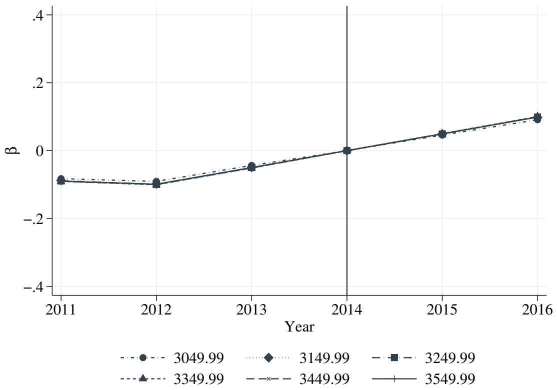
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

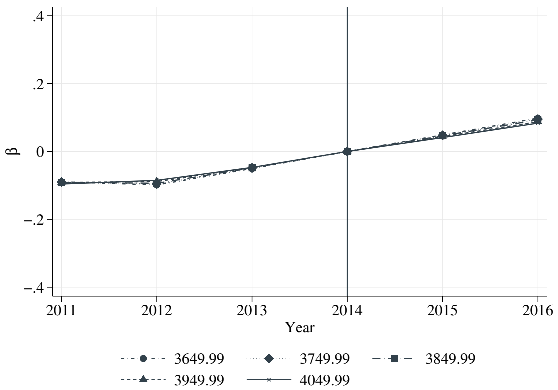
Figure A.80: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 3 (augmented industries). No firm controls.



(a) Thresholds: 2,449.99 to 2,949.99



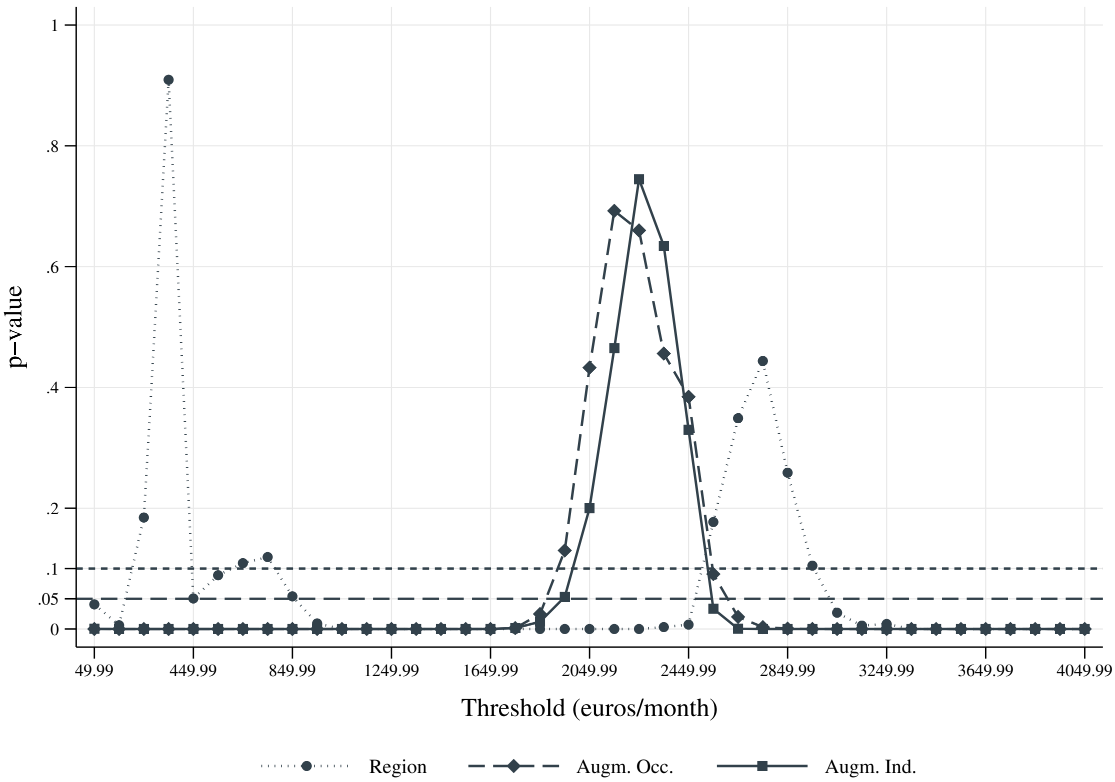
(b) Thresholds: 3,049.99 to 3,549.99



(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.81: P-values of joint significance – Monthly earnings. No firm controls.

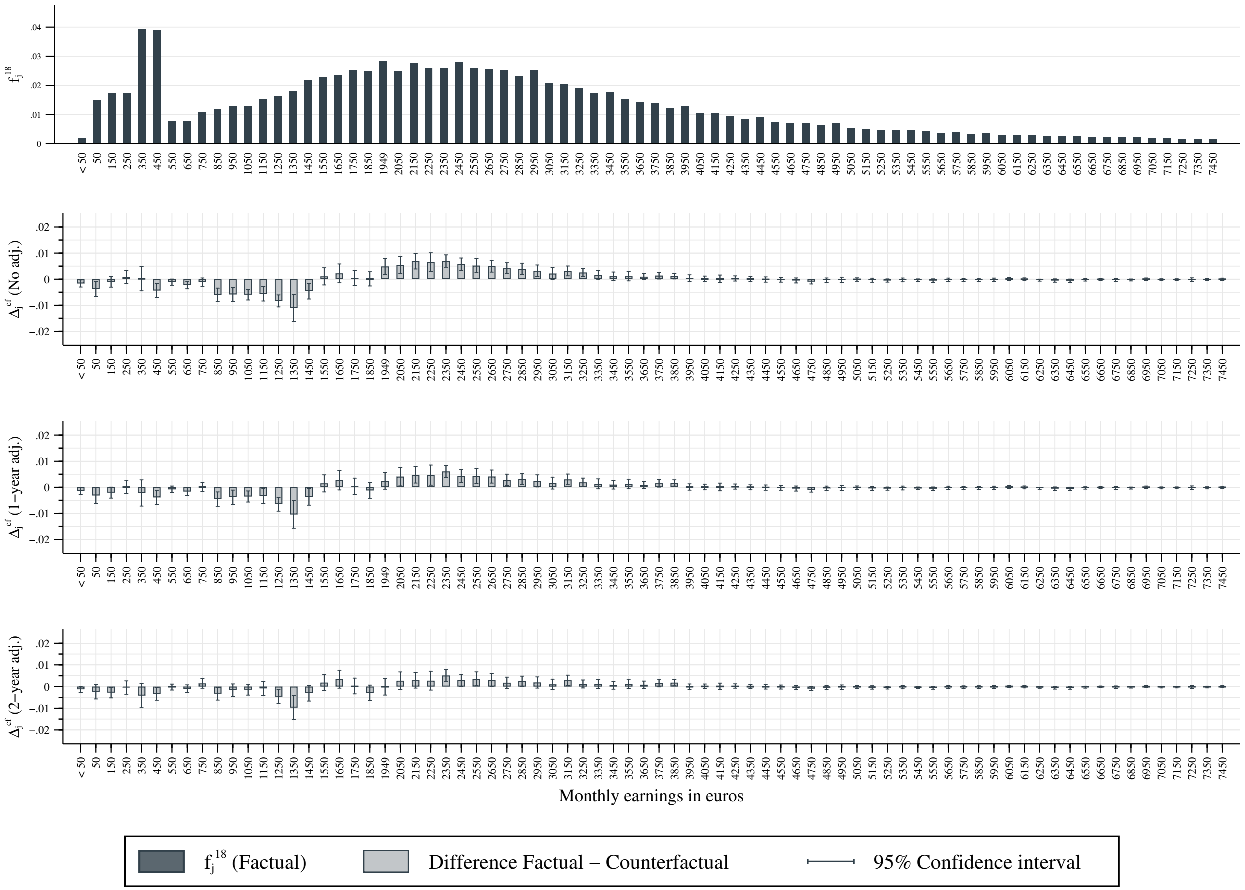


Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

No controls

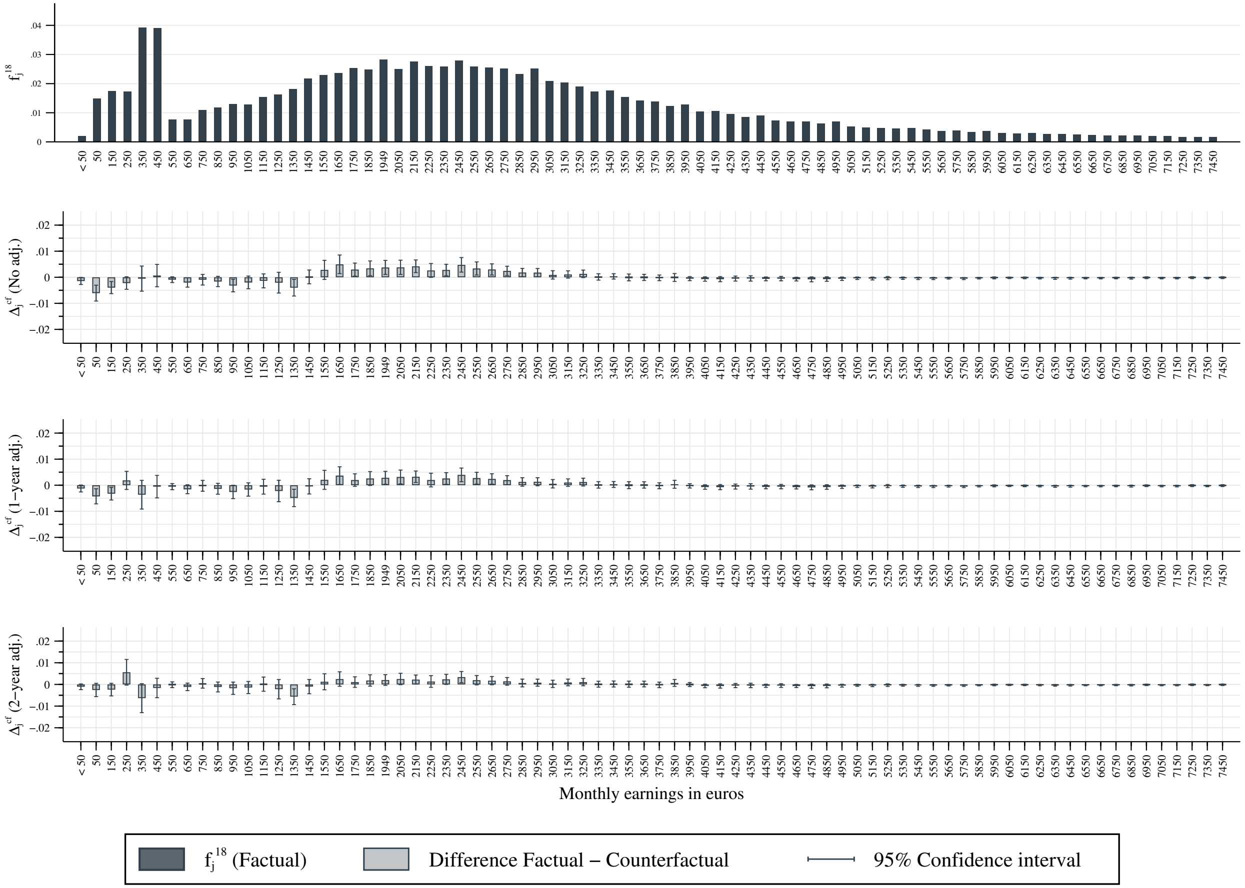
— Intentionally left blank —

Figure A.82: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 1: Regions. No controls.



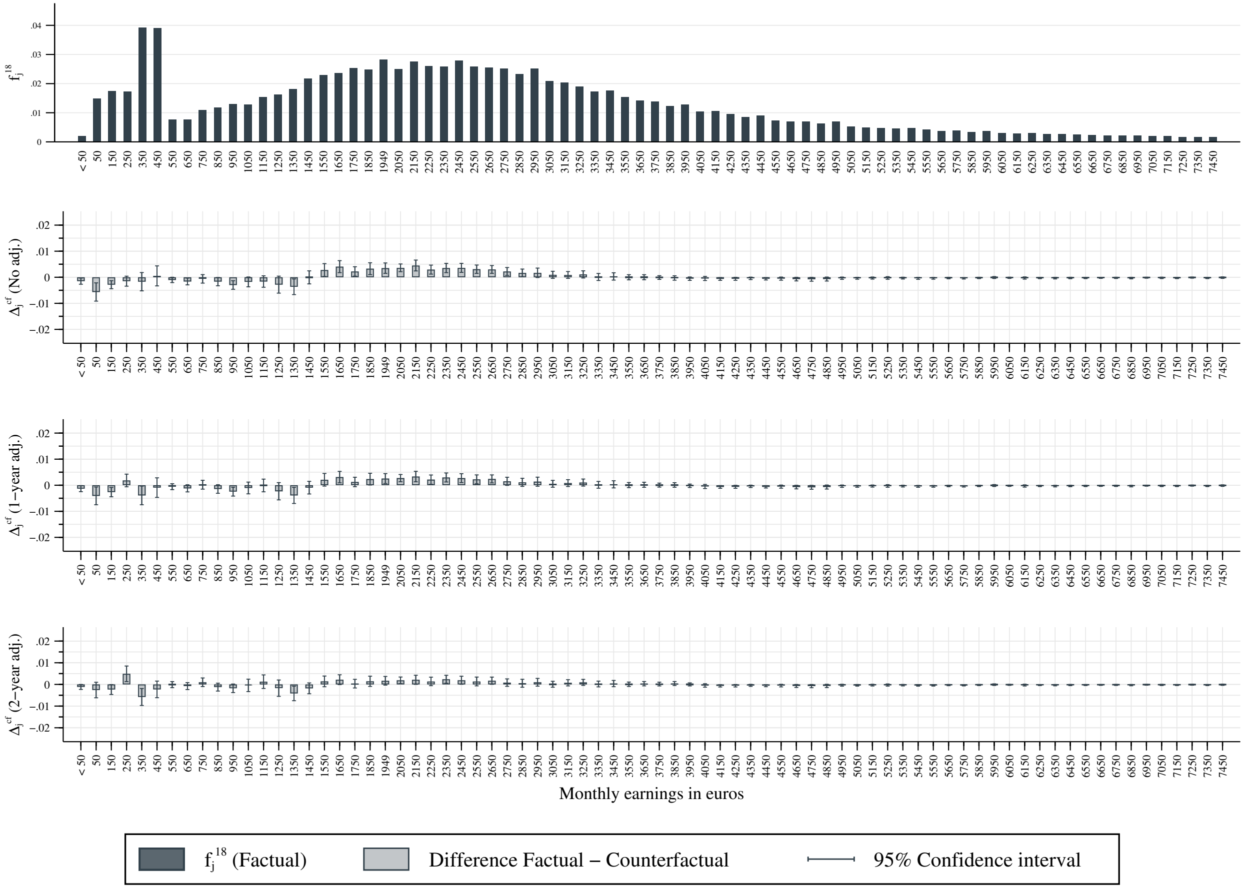
Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.83: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 2: Augmented occupations. No controls.



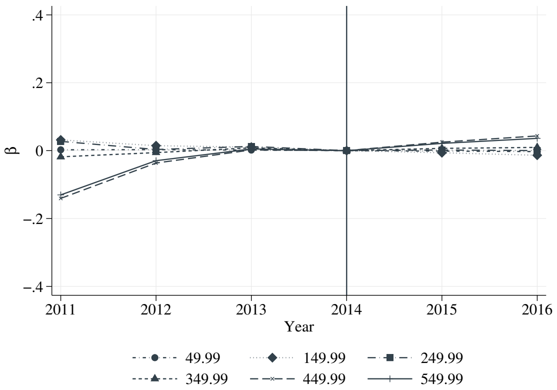
Notes: The x-axis shows monthly wage bins. For example, the ‘1050’ bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

Figure A.84: 2018 Factual monthly earnings distribution and treatment effects due to minimum wage for different trend adjustment regimes. Bite 3: Augmented industries. No controls.

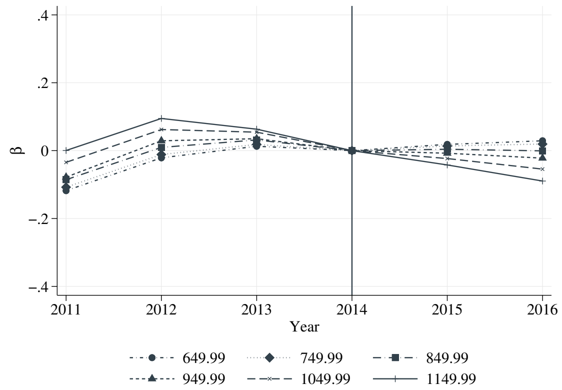


Notes: The x-axis shows monthly wage bins. For example, the '1050' bin comprises monthly earnings in the interval [1,050; 1,149] euros. The first row depicts the bin-wise factual distributional mass in 2018. The three lower panels show the isolated minimum wage effect under different pre-trend adjustment regimes. 95% bootstrap confidence intervals (100 replications, clustered at treatment level).
Source: GSES 2014/18, DGUV-IAB 2011-14, own calculations.

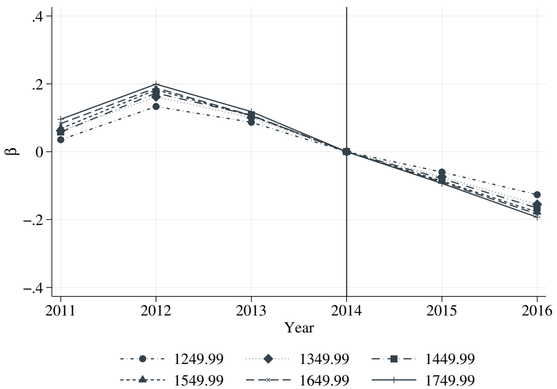
Figure A.85: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
– Monthly earnings (lower to middle thresholds), bite 1 (region). No controls.



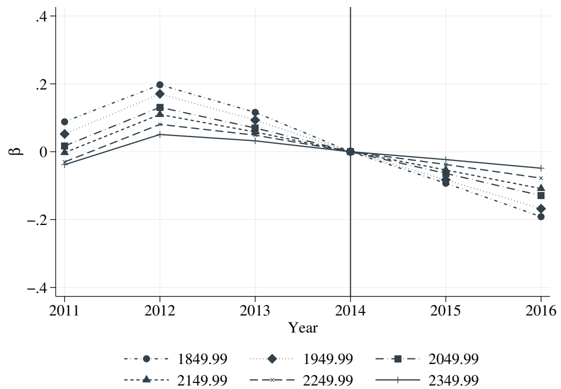
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



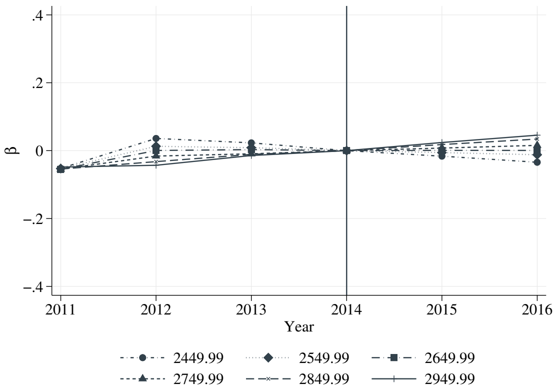
(c) Thresholds: 1,249.99 to 1,749.99



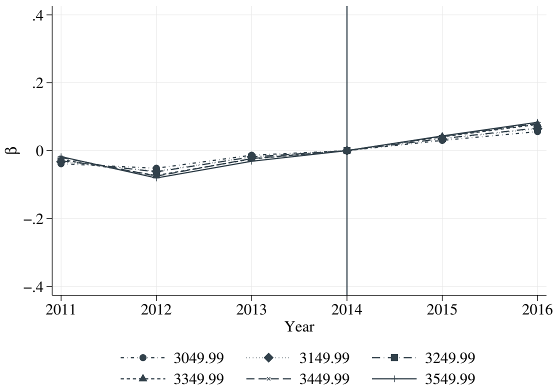
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

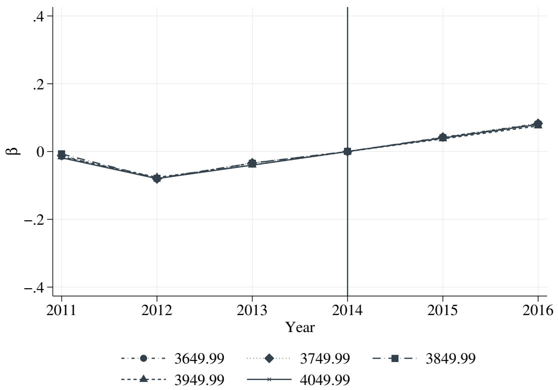
Figure A.86: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 1 (region). No controls.



(a) Thresholds: 2,449.99 to 2,949.99



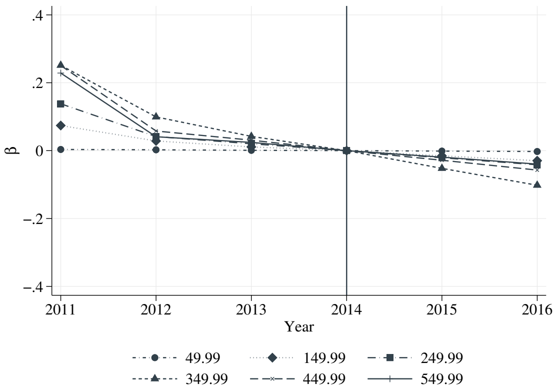
(b) Thresholds: 3,049.99 to 3,549.99



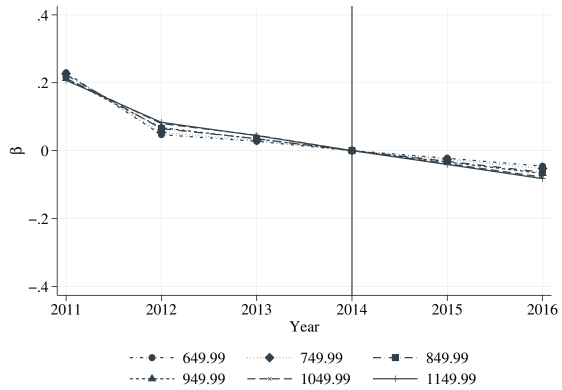
(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

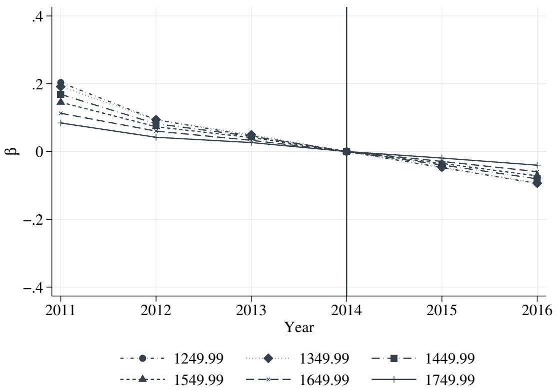
Figure A.87: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data
– Monthly earnings (lower to middle thresholds), bite 2 (augmented occupations). No controls.



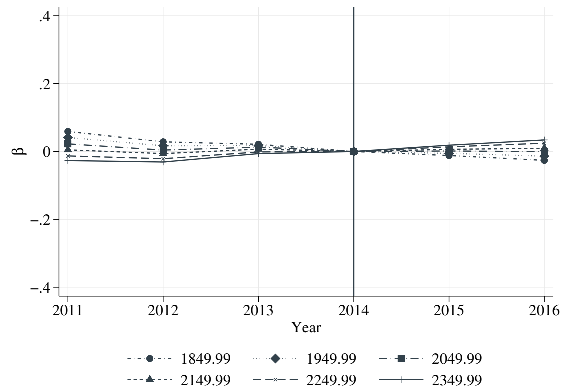
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1,149.99



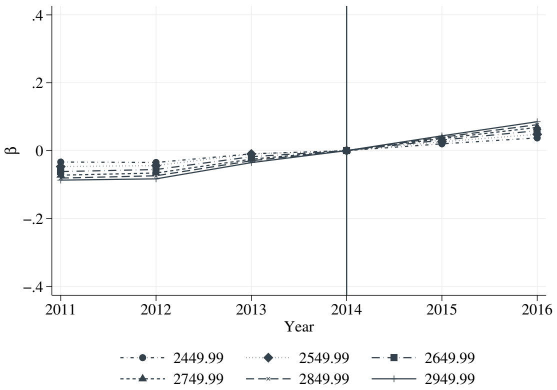
(c) Thresholds: 1,249.99 to 1,749.99



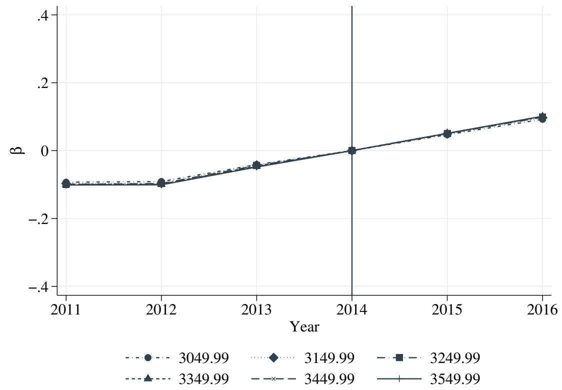
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

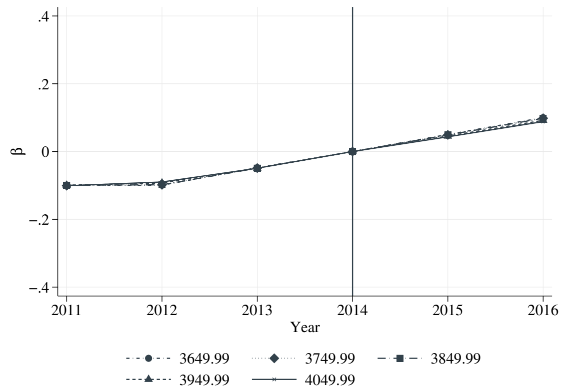
Figure A.88: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 2 (augmented occupations). No controls.



(a) Thresholds: 2,449.99 to 2,949.99



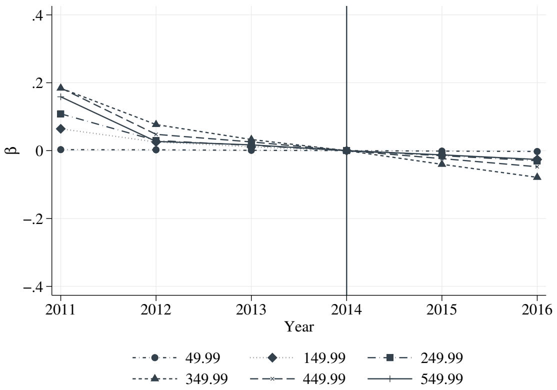
(b) Thresholds: 3,049.99 to 3,549.99



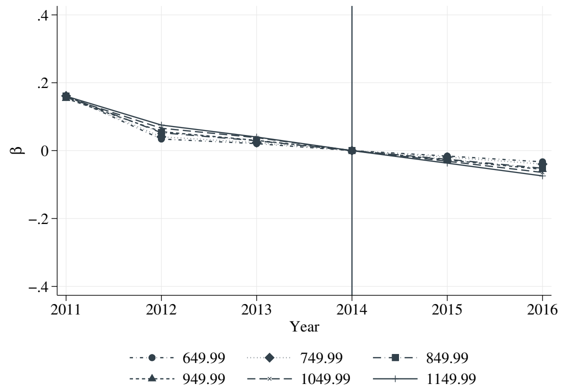
(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

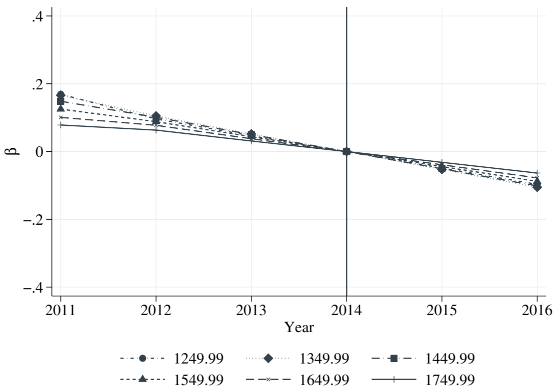
Figure A.89: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (lower to middle thresholds), bite 3 (augmented industries). No controls.



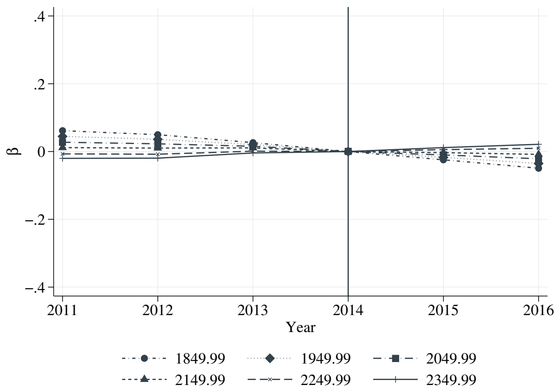
(a) Thresholds: 49.99 to 549.99



(b) Thresholds: 649.99 to 1149.99



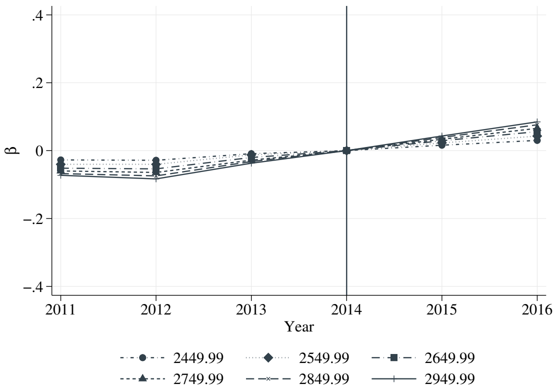
(c) Thresholds: 1,249.99 to 1,749.99



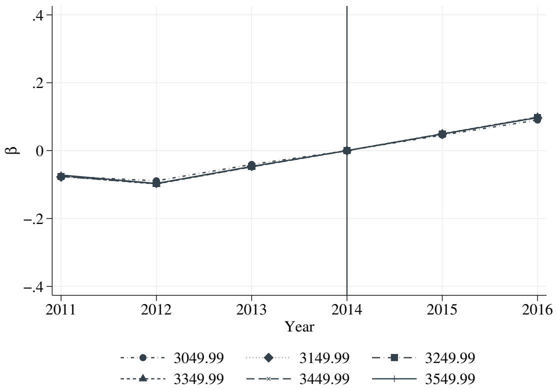
(d) Thresholds: 1,849.99 to 2,349.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

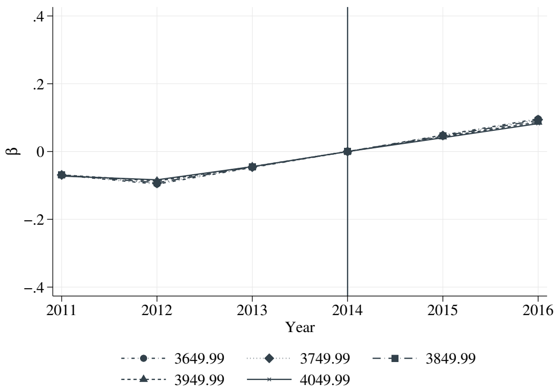
Figure A.90: Pre-treatment estimates of treatment coefficients using *DGUV-IAB* data – Monthly earnings (upper thresholds), bite 3 (augmented industries). No controls.



(a) Thresholds: 2,449.99 to 2,949.99



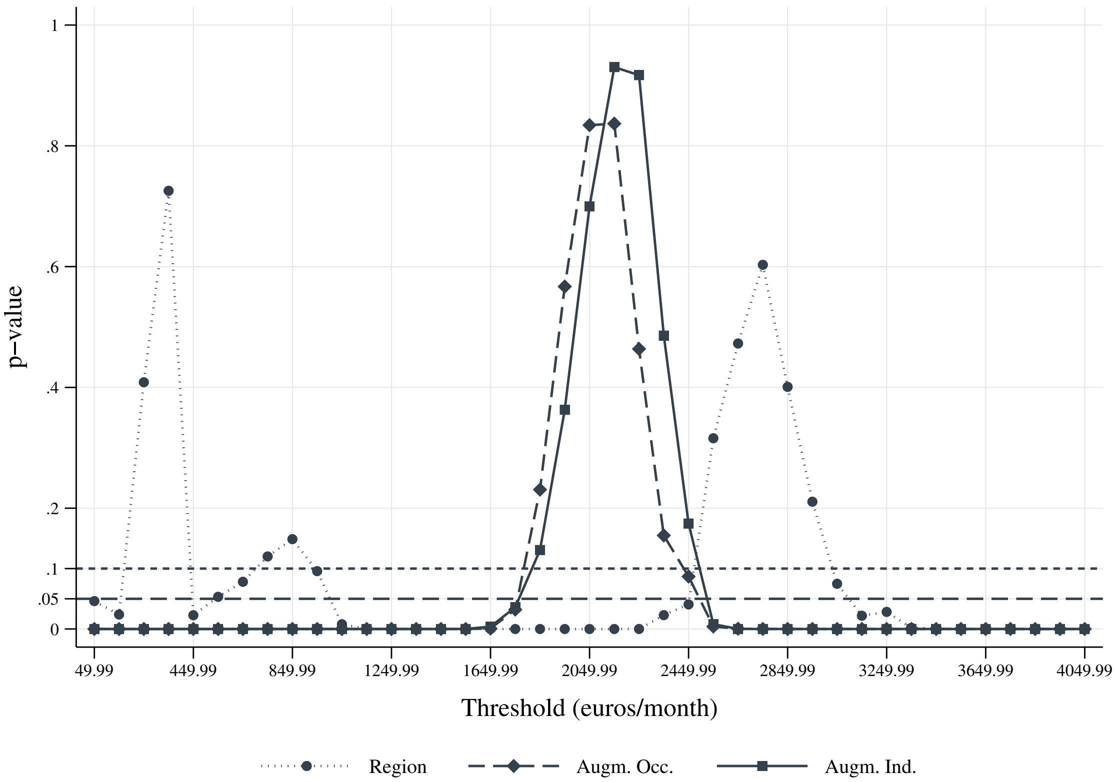
(b) Thresholds: 3,049.99 to 3,549.99



(c) Thresholds: 3,649.99 to 4,049.99

Notes: Estimates for the treatment effect, $\hat{\beta}_z^t$, in the pre-treatment periods 2011-2014 as specified in (2.11). Base period: 2014. Values in 2015 and 2016 refer to linearly extrapolated trends using the estimates from 2012, 2013, and 2014.
Source: DGUV-IAB 2011-14, own calculations.

Figure A.91: P-values of joint significance – Monthly earnings. No controls.



Notes: Values indicate p-values from a Wald test for joint significance of all relevant pre-trend estimates $H_0: \beta_z^{2011} = \beta_z^{2012} = \beta_z^{2013} = 0$ in (2.11) for a given threshold z based on bootstrap (100 replications, clustered at treatment level).
Source: DGUV-IAB 2011-14, own calculations.

Chapter 3

Combining Difference-in-Differences and Recentered Influence Functions to Estimate Quantile Treatment Effects – Pitfalls and Remedies*

*This chapter is based on Rümmele, M. (2024). Combining Difference-in-Differences and Recentered Influence Functions to Estimate Quantile Treatment Effects – Pitfalls and Remedies. *Unpublished manuscript, University of Tübingen*.

3.1 Introduction

The difference-in-differences (DiD) approach remains one of the most widely employed approaches when it comes to evaluating policy effects such as changes in minimum wage policies or tax regimes. It is fair to say that, despite its long-standing tradition in the literature, DiD techniques have not forfeited their topicality but have in fact experienced a renaissance. Numerous recent contributions examine traditional DiD approaches more formally, explore distributional extensions of DiD approaches, or hint at potential shortcomings in specific research settings (see the review in Roth et al., 2023).

Most contributions focus on the effect of policies on the mean value of an outcome by leveraging a common trend assumption for the mean. While focusing on the center of the distribution is often reasonable, some policies are explicitly designed to affect certain parts of the outcome distribution. Therefore, employing suitable approaches to estimate such effects is of obvious interest in applied empirical work. Assessing distributional effects has, for example, been of interest in evaluations of minimum wage policies, as examined in Biewen et al. (2022), Bossler and Schank (2023), and Cengiz et al. (2019), or in the context of tax credits targeting individuals at the bottom of the income distribution (e.g., Hoynes and Patel, 2018).

While there are numerous contributions that consider the identification and estimation of distributional or quantile treatment effects in cross-sectional settings (e.g., Firpo, 2007; Rothe, 2010, 2012; Frölich and Melly, 2013; Firpo and Pinto, 2016; Powell, 2020), there are plenty of recent contributions that explicitly examine distributional treatment effects in a DiD setting as well (e.g., Kim and Wooldridge, 2024; Roth and Sant'Anna, 2023; Callaway and Li, 2019; Callaway et al., 2018; Fan and Yu, 2012; Bonhomme and Sauder, 2011; Athey and Imbens, 2006). There is, however, no unified methodological toolkit to estimate these distributional treatment effects. If one aims at estimating distributional treatment effects using a regression approach, conditional quantile regression techniques (CQR, Koenker and Bassett Jr., 1978) are a natural starting point. However, a CQR approach is restrictive if one is interested in unconditional quantile treatment effects and if one wants to control for a vector of covariates within a linear regression. Since, in many cases, the target of interest is not the effect for a specific group but, more generally, the effect along the marginal distribution, combining a DiD rationale with *unconditional* quantile regressions based on recentered influence functions (RIF, Firpo et al., 2009) might, at first sight, come across as a compelling alternative. As a matter of fact, recent contributions have applied RIF approaches to estimate distributional treatment effects within a DiD framework (Havnes and Mogstad, 2015; Gregory and Zierahn, 2022; Caliendo et al., 2023; Bossler and Schank, 2023; Bossler et al., 2024; Pusch, 2024, reviewed in greater detail below). As these articles focus on applying the RIF as a mere tool to estimate quantile treatment effects, a formal discussion regarding potential issues or shortcomings of the RIF in a DiD setting is either very brief or not present at all.

Hence, the first aim of this paper is to extend the literature on potential shortcomings of RIF approaches that stem from 'technical' issues in certain settings beyond the cross-sectional case. Although RIF approaches are conveniently implementable and allow for the consideration of unconditional rather than conditional quantiles, it is often overlooked that they are not inherently designed for treatment effect analyses involving discrete group comparisons. In fact, RIF regressions are, by construction, approximations and have proved to be particularly fruitful in the context of decomposing changes in inequality measures into composition and wage structure effects (Firpo et al., 2018; Fortin et al., 2011; for recent applications see, e.g., Biewen and Seckler, 2019; Baumgarten et al., 2020).

The disadvantages of RIF approaches in cross-sectional settings that involve a discrete comparison of sub-populations have already been discussed in previous contributions. Borgen et al. (2022) discuss misconceptions underlying the use of a RIF approach to estimate quantile treatment effects in a cross-sectional setting, noting that several earlier contributions have misleadingly employed RIF approaches to estimate a quantile treatment effect of some sort. In the same vein, Rios-Avila and Maroto (2024) discuss the potential of using an adjusted RIF regression approach to estimate quantile treatment effects. Their approach is based on the notion that the RIF is merely a representation of an unconditional distributional statistic. In turn, this can be used to define a compound RIF that includes 'updated' RIFs for both the treated and untreated groups, which can then serve as the new dependent variable in a linear regression. This approach has also been briefly described in Rios-Avila (2020) within the broader context of using RIFs to estimate 'inequality treatment effects' (Firpo and Pinto, 2016). While these contributions discuss the potential and misuse of RIFs for treatment effect analyses in a cross-sectional setting, this paper specifically aims to contribute to this strand of literature by adding insights on the possibility of using RIF regression techniques within a DiD framework.

To the best of our knowledge, the contribution by Kottelenberg and Lehrer (2017) is the only instance that contains an explicit critique of a RIF-DiD approach (Kottelenberg and Lehrer, 2017, p. 623, footnote 24). However, they neither examine potential solutions nor provide a formal representation of the issues they identify. Hence, this study is the first to thoroughly examine the shortcomings of combining a RIF approach with a DiD framework, discuss ways to overcome them, and link suitable RIF approaches to the necessary identifying assumptions.

In particular, it is shown that a simple combination of a RIF and a DiD approach ultimately results in a 'pooled' RIF approach that suffers from structural shortcomings. Crucially, a pooled RIF-DiD approach does not formally allow for the estimation of the quantile treatment effect on the treated (QTT), as effects are approximated around misleading points of the outcome distribution. To address this issue, two approaches are proposed that utilize certain RIF 'mechanics'. Both approaches are connected to the identifying assumptions required by distributional DiD methods discussed in prior studies. The first approach leverages the principle that a RIF method is based on first-order approximations around, for example, a specific initial quantile of interest. It is demonstrated that a RIF approach, in which effects are approximated around the correct quantile of interest, can, in principle, approximate the true QTT, whereas approximative errors persist. Perhaps somewhat surprisingly, and in contrast to the formulated critique by Kottelenberg and Lehrer (2017), it is shown that a 'correct' RIF approximation for the QTT does not require a common trend in *quantiles*, but rather a common trend in *population shares*. As a second approach, it is also possible to formulate a RIF-DiD framework that is built upon a common trend in quantiles. This approach revolves around the insight that RIFs provide a means of expressing unconditional quantiles, thus allowing for the formulation of a compound RIF that suitably combines the four involved sub-populations in a DiD context. This approach mirrors the method proposed by Rios-Avila and Maroto (2024) for the cross-sectional case. In addition to an analytic treatment, the appropriateness of these approaches is studied in the context of a Monte Carlo simulation.

As a second contribution, this paper more generally aims at adding to a growing strand of literature that discusses unintended estimation outcomes of well established econometric techniques, especially in the context of estimating treatment effects. For example, Słoczyński (2022) and De Chaisemartin and d'Haultfœuille (2020) both show that ignoring heterogeneity in treatment effects can be problematic as the implied weights dilute the actually estimated treatment effects when using canonical econometric techniques, i.e., OLS in the former and two way fixed effects (TWFE) in the latter contribution. Goodman-Bacon (2021) also shows that the canonical TWFE DiD estimator potentially provides biased estimates if more than two time periods are involved and individuals are not all treated at the same point in time. All these contributions follow the principle of clearly distinguishing between a precise formulation of the treatment effect that is to be *identified*, and an *estimation* and inference approach that is appropriate for the given case. By the same token, this paper aims at clearly distinguishing between the object that is to be *identified* and a potential *estimation* technique that could be employed to that end. In particular, it is argued that the simple combination of a RIF and a DiD approach provides deceptive estimates of a distributional treatment effect since it is based upon a misleading formulation of the object that is to be estimated. Concretely, both the formal examination and the Monte Carlo study results highlight that the pooled RIF approach structurally fails to provide a sensible approximation for the QTT.

The remainder of the paper is structured as follows. Section 3.2 introduces the RIF and generally outlines the particularities of the approach. Section 3.3 introduces the distributional DiD case in the most simplistic 2×2 scenario and the underlying assumptions to identify distributional treatment effects. Section 3.4 formally examines different RIF approaches regarding their appropriateness for estimating the QTT within a simple DiD framework. In section 3.5, a Monte Carlo simulation study is conducted to compare the RIF approaches. Section 3.6 reexamines a distributional DiD approach related to the expansion of the Earned Income Tax Credit (EITC) in the US, as discussed by Hoynes and Patel (2018), to compare the three approaches in a 'real-world' application. Finally, section 3.7 concludes.

3.2 Recentered influence functions

In their seminal contribution, Firpo et al. (2009) suggest the use of (recentered) influence functions to examine the effect of small covariate-induced shifts on unconditional quantiles. The advantage of this approach is that it does not require a global inversion of an implied counterfactual distribution if one is solely interested in the effect around the τth quantile of some outcome distribution. Rather, a RIF approach in this case entails a local approximation of the effect around the initial quantile of interest. In the following, the RIF approach as introduced in Firpo et al. (2009) is briefly described.

At the heart of the RIF approach is the concept of *influence functions* (Hampel, 1974). Formally, influence functions are a special form of Gateaux derivatives that describe the effect of a small perturbation in the distribution F_Y on a distributional statistic $v(\cdot)$.[1] Intuitively speaking, influence functions are a vehicle through which changes of distributional statistics due to marginal shifts in the original distribution can be examined. One could figuratively think of influence functions as providing a measure of the impact induced by adding or removing an observation $y \in \mathscr{Y} \subset \mathbb{R}$ on a distributional statistic $v(F_Y(y))$. Formally, following the notion of being a directional *derivative*, influence functions refer to the infinitesimally small shift from some initial distribution, F_Y, towards Δ_Y, with Δ_Y being a point mass. This reads as

$$IF(y;v,F_Y) \equiv \lim_{t\to 0} \frac{v(F_{(1-t)F_Y+t\Delta_Y}) - v(F_Y)}{t},$$

with $F_{(1-t)F_Y+t\Delta_Y}$ being a mixture distribution that is shifted towards Δ_Y by magnitude of t. Reminiscent of Taylor approximations, this directional derivative can be used to formulate first-order approximations around the initial distributional statistic. Specifically, an approximation of the distributional statistic evaluated at a distribution that is shifted towards G_Y but remains in the proximity of F_Y can be expressed as

$$v(F_{(1-t)F_Y+t\cdot G_Y}) \approx v(F_Y) + t \cdot \int IF(y;v,F_Y)\, d(G_Y - F_Y)(y).$$

Considering the case with $t = 1$ and $G_Y = \Delta_Y$, Firpo et al. (2009) coin the leading term of such a first order approximation the *recentered* influence function, since, in this case, the leading term reads[2]

$$RIF(y;v,F_Y) = v(F_Y) + \int IF(y;v,F_Y)\, d\Delta_Y(y) = v(F_Y) + IF(y;v,F_Y). \tag{3.1}$$

Using the fact that the expected value of the influence function is 0, the expected value of

[1] See for example the outline in Hampel et al. (1986).

[2] To see why this is the case, not that, by virtue of being a point mass, it holds that $\int a(y)\, d\Delta_Y(y) = a(y)$.

the RIF as defined in (3.1) yields the distributional statistic of interest again:[3]

$$v(F_Y) = \mathbb{E}[RIF(Y;v,F_Y)] = \int RIF(y;v,F_Y)\, dF_Y(y). \tag{3.2}$$

This conveniently allows for modeling perturbations in the outcome distribution in terms of shifts in some covariates by virtue of the law of iterated expectations:

$$v(F_Y) = \iint RIF(y;v,F_Y)\, dF_{Y\,|\,X}(y\,|\,X=x)\, dF_X(x) = \int \mathbb{E}[RIF(Y;v,F_Y)\,|\,X=x]\, dF_X(x),$$

which can be modeled using a linear specification:

$$v(F_Y) = \mathbb{E}\left[\mathbb{E}[RIF(Y;v,F_Y)\,|\,X]\right] = \mathbb{E}[X]\beta_{v(\cdot)}, \tag{3.3}$$

with $\mathbb{E}[X]$ ($\beta_{v(\cdot)}$) being a K-dimensional row (column) vector including a constant. Hence, the effect of a ceteris paribus location shift in one of the covariates, X_k, is modeled as $\beta_{v(\cdot),k}$, which can be estimated by OLS.

Unconditional quantile regression using RIF-OLS

Below, DiD treatment effects for unconditional *quantiles* are to be considered. Based on the general formulation in (3.1), the RIF for the τth unconditional quantile of an outcome variable y is given by

$$RIF(y;q_\tau,F_y) = q_\tau + \frac{\tau - \mathbb{1}\left[y \leq q_\tau\right]}{f_Y\left(q_\tau\right)}, \tag{3.4}$$

with $f_Y(q_\tau)$ being the probability density function (pdf) evaluated around the initial unconditional quantile of interest, q_τ. To shed more light on the particularities involved in the interpretation of RIF-OLS coefficients, the RIF regression coefficients for the case of unconditional quantiles are examined in greater detail. Reconsidering the term in (3.4) and assuming linearity of the conditional expected value function, one can write the expected value, conditional on a vector of covariates X, as

$$\mathbb{E}[RIF(Y;q_\tau,F_Y)\,|\,X] = q_\tau + \frac{\tau - \mathbb{E}[\mathbb{1}[Y \leq q_\tau]\,|\,X]}{f_Y(q_\tau)} = X\beta_\tau, \tag{3.5}$$

[3]Intuitively, this arises because the influence function is defined to capture deviations from the functional, $v(\cdot)$, and thus, integrating these deviations over F_Y cancels them out, resulting in $\mathbb{E}[IF(Y;v,F_Y)] = \int IF(y;v,F_Y)\, dF_Y(y) = 0$.

with X (β_τ) being a K-dimensional row (column) vector as above. In most applications, the involved vector of coefficients, β_τ, derive their interpretation from the location-shift thought experiment of a given covariate X_k, i.e., the examination of the impact of a small shift, $\mathbb{E}[X_k] \to \mathbb{E}[\widetilde{X}_k]$, on the τth unconditional quantile. As noted above, considering the expected value of (3.5) yields the unconditional quantile, i.e.,

$$q_\tau = \mathbb{E}[\mathbb{E}[RIF(Y;q_\tau,F_Y)\,|\,X]] = q_\tau + \frac{\tau - \mathbb{E}[\mathbb{P}[Y \le q_\tau\,|\,X]]}{f_Y(q_\tau)} = \mathbb{E}[X]\beta_\tau.$$

This linear representation allows for approximating changes in q_τ implied by a small location shift in X using linear regression techniques. To see this, note that under the linearity assumption, the term in the numerator, $\mathbb{P}[Y \le q_\tau\,|\,X]$ can be expressed as

$$\mathbb{P}[Y \le q_\tau\,|\,X_{-k}, X_k] = \gamma_{-k,\tau}X_{-k} + \gamma_{k,\tau}X_k, \tag{3.6}$$

with $\gamma_{-k,\tau}$ (X_{-k}) being a row (column) vector of dimension $K-1$, not involving the kth element, and $\gamma_{k,\tau}$ as well as X_k are scalars referring to the kth element, respectively. Using this, one can illustrate the underlying mechanisms of a RIF approximation around q_τ. Denoting $\widetilde{q}_\tau$ as the actual counterfactual quantile under $\mathbb{E}[\widetilde{X}_k]$, the difference $\widetilde{q}_\tau - q_\tau$ can be linearly approximated as

$$\begin{aligned}\Delta q_\tau = \widetilde{q}_\tau - q_\tau &\approx \mathbb{E}[\mathbb{E}[RIF(\cdot)\,|\,\widetilde{X}_k, X_{-k}]] - \mathbb{E}[\mathbb{E}[RIF(\cdot)\,|\,X_k, X_{-k}]]\\ &= \left(q_\tau + \frac{\tau - \mathbb{E}[\mathbb{P}[Y \le q_\tau\,|\,\widetilde{X}_k, X_{-k}]]}{f_Y(q_\tau)}\right) - \left(q_\tau + \frac{\tau - \mathbb{E}[\mathbb{P}[Y \le q_\tau\,|\,X_k, X_{-k}]]}{f_Y(q_\tau)}\right).\end{aligned}$$

$$\implies \Delta q_\tau \approx -\frac{\overbrace{\gamma_{k,\tau}(\mathbb{E}[\widetilde{X}_k] - \mathbb{E}[X_k])}^{=\Delta\tau}}{f_Y(q_\tau)} = \beta_{k,\tau}(\mathbb{E}[\widetilde{X}_k] - \mathbb{E}[X_k]), \tag{3.7}$$

where $\beta_{k,\tau}$ is the kth coefficient of a regression involving the RIF as the dependent variable, and $\gamma_{k,\tau}$ refers to the kth coefficient obtained from a linear regression modeling the probability of falling below q_τ, as in (3.6). As shown in (3.7), the implied changes in the population share, denoted by $\Delta\tau$, can be thought of as being induced by the shift $\mathbb{E}[X_k] \to \mathbb{E}[\widetilde{X}_k]$. The most common way to estimate the effect of a small shift in the kth covariate on the τth quantile is to compute the RIF values for the τth quantile according to (3.4) and linearly project it on a vector of covariates and a constant, i.e., directly estimating $\beta_{k,\tau}$ in (3.7). However, the specification outlined in (3.7) suggests an alternative two-step procedure that yields the same coefficient as directly estimating $\beta_{k,\tau}$.

Firstly, estimate coefficients in (3.6) using the following estimation equation:

$$\mathbb{1}[Y_i \leq \widehat{q}_\tau] = \gamma_{-k,\tau} X_{i,-k} + \gamma_{k,\tau} X_{i,k} + \varepsilon_i,$$

with $\widehat{q}_\tau$ being an estimator of the unconditional τth quantile. As described in Koenker and Bassett Jr. (1978), the quantile estimate can be obtained by means of $\widehat{q}_\tau = \arg\min_u \sum_i (\tau - \mathbb{1}[Y_i - u \leq 0]) \cdot (Y_i - u)$. Secondly, translate these estimated effects on the propensity to fall below q_τ into *quantile* effects by dividing by the slope of the cumulative distribution function (cdf) evaluated at the initial quantile of interest:

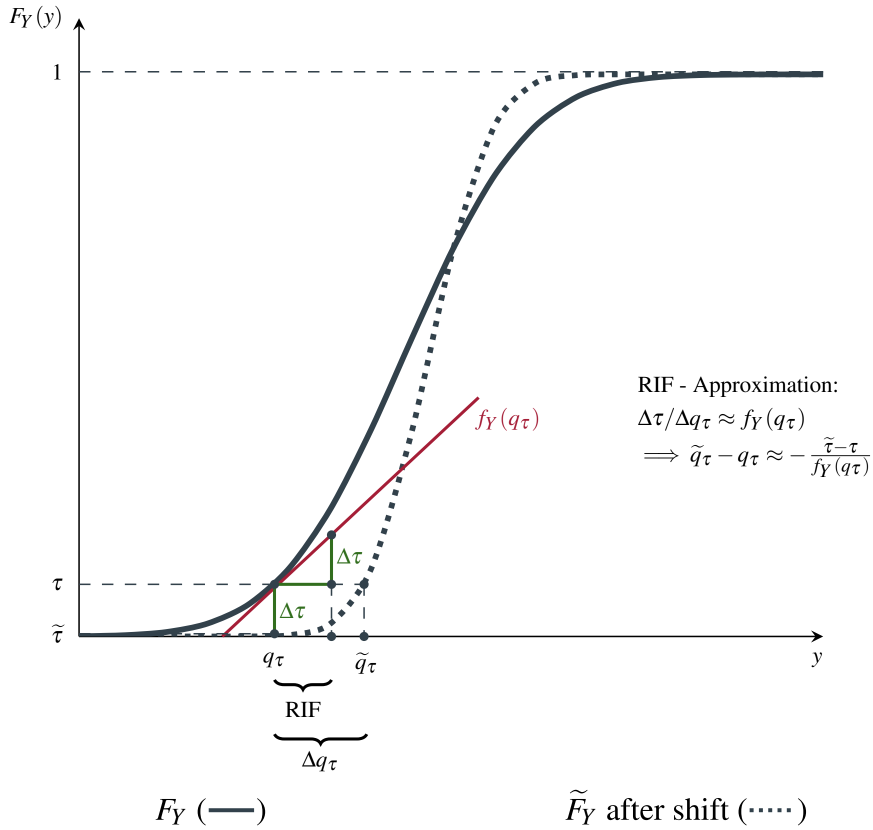
$$\widehat{\beta}_{k,\tau} = -\frac{\widehat{\gamma}_{k,\tau}}{\widehat{f}_Y(\widehat{q}_\tau)}, \tag{3.8}$$

with $\widehat{f}_Y(\widehat{q}_\tau)$ being a kernel density estimator for the distributional mass around the given initial quantile of interest (see Firpo et al., 2009, for further details). Summarized, a RIF regression for quantiles essentially contains a 'distribution regression' (Chernozhukov et al., 2013) part that models the effect on the population share ($\Delta\tau$), as well as a 'translation part', by which the $\Delta\tau$ is translated into the quantile effect (Δq_τ).

The underlying intuition of a 'distribution regression' and a 'translation' part being involved in any RIF regression for unconditional quantiles can be graphically illustrated as in figure 3.1.[4] Beside displaying the translation idea, figure 3.1 visualizes the notion that RIF-based techniques only provide first-order approximations of the actual quantile effect. In summary, the RIF regression approach allows researchers to conveniently model small (covariate-induced) shifts in distributional statistics, such as unconditional quantiles, using linear regression techniques. However, this simplicity comes with the limitation that the effects are only locally valid in the vicinity of the initial distribution.

[4]Comparable graphical intuitions can, for example, be found in Fortin et al. (2011).

Figure 3.1: Intuition for RIF-OLS coefficients in the case of unconditional quantiles



Notes: Figure displays graphical representation of the notion underlying a RIF approximation in the context of unconditional quantiles. Both the actual (true) effect, Δq_τ, and the RIF approximation are illustrated. The implied shift in population shares, $\Delta\tau = \widetilde{\tau} - \tau$, can be obtained by estimating parameters of a distribution regression model as in (3.6). The RIF approximation is intuitively obtained by means of the triangular formula, $\Delta q_\tau \approx -(\Delta\tau/f_Y(q_\tau))$, which operationalizes the translation of a vertical into a horizontal shift. The implied $\widetilde{\tau}$ can be thought of as arising due to covariate-induced shifts, as indicated in (3.7).

RIF-OLS: Particularities

In cases other than the mean, the RIF-OLS approach is built upon certain RIF-inherent assumptions. Concretely, it needs to be taken into account that a new dependent variable has to be generated. The latter turns out to be restrictive in certain settings, and particularly so in a RIF-DiD framework. Technically speaking, problems may arise due to *misspecifications on the left-hand side*. Specifically, in cases other than the mean, the RIF requires certain parameters that essentially define the locality of the RIF. Reconsidering the RIF for the case of quantiles as in (3.4), it is evident that both an estimator for the quantile of interest, $\widehat{q}_\tau$, and a density estimator evaluated at $\widehat{q}_\tau$, $\widehat{f}_Y(\widehat{q}_\tau)$, are needed up front. Most importantly, these elements are specific to the distribution of Y, i.e., one needs to take into account which distribution the quantiles actually refer to. In cases where the empirical approach involves several sub-populations to be discretely compared with one another, as in a DiD framework, it is essential to be explicit about these up-front elements and the specific distributions they are based on.

Another crucial particularity of RIF-based approaches is related to potentially poor approximations of the actual effect. If the shift implies a counterfactual distribution that is shaped differently than the original distribution and/or the implied effect is very large, the RIF effect becomes a poor approximation of the actual effect since RIF-OLS coefficients merely provide a first-order approximation of the true effect as illustrated in figure 3.1. Intuitively, if the original and the implied distribution were both uniformly distributed with the same slope (i.e., a mere location shift effect) the approximation would be perfect. In the context of RIFs for unconditional quantiles, another source of poor approximations manifests itself through the RIF containing a kernel density in its denominator. This implies additional complications that are related to the stability of the estimated effect if the outcome distribution is characterized by bunching and discrete mass points around the initial quantile of interest. The latter is, for example, of particular importance in the context of evaluations of a minimum wage introduction using a DiD framework due to the implied mass point around the minimum wage threshold (Biewen et al., 2022).

3.3 DiD and distributional treatment effects

In the following, two different concepts for characterizing distributional treatment effects are presented. Moreover, the respective identifying assumptions are outlined. In doing so, we focus on the simplest 2×2 form of DiD analyses. That being said, issues such as treatment effect heterogeneity, as discussed in De Chaisemartin and d'Haultfœuille (2020) and Goodman-Bacon (2021), are deliberately abstracted from to concentrate on the RIF-inherent components involved in the RIF-DiD approaches examined below.

The identifying assumptions for the 2×2 DiD case are best understood in terms of a potential outcome framework. In what follows, the identification of distributional treatment effects will be established such that repeated cross-sectional sampling is accommodated as well.[5] Denote treatment group membership by $D_i = d \in \{0,1\}$, with 0 indicating membership of the control and 1 indicating membership of the treatment group, respectively. By the same token, denote the pre- and post-treatment period by $T_i = t \in \{0,1\}$. Suppose further that one can obtain iid draws from the distributions $(Y_i, D_i) \mid T_i = 0$ and $(Y_i, D_i) \mid T_i = 1$. Henceforth, these iid draws from group $D_i = d$ at point in time $T_i = t$ will be referred to as $Y_i(d,t)$.

Ultimately, the treatment effect is defined by comparing a factual with a counterfactual (potential) outcome absent treatment in the post-treatment period. Denote potential outcomes in the absence of treatment for a given individual i in group $D_i = d$ at point in time $T_i = t$ as $Y_i^0(d,t)$. Under the assumption of no anticipatory effects in the pre-treatment period, all potential outcomes in the pre-treatment period are observable since treatment occurs only in $t = 1$, i.e., $Y_i^0(d,0) = Y_i(d,0)$ for $d \in \{0,1\}$. In the post-treatment period, however, only one state of the world for either group is observable, encompassing the stable unit treatment value assumption (SUTVA), i.e., the treatment status of one individual must not depend on the treatment status of another individual. Consequently, while $Y_i^0(0,1) = Y_i(0,1)$ is observable, the potential outcome for treated individuals in the post-treatment period, $Y_i^0(1,1)$, is unobservable and therefore only identifiable under additional assumptions. These assumptions differ depending on the statistic of interest.

Identifying distributional treatment effects within a DiD framework

Essentially, there are two ways of defining treatment effects along the entire support of the outcome variable, the *distributional treatment effect on the treated (DTT)* as well as the *quantile treatment effect on the treated (QTT)*. Below, a formal definition as well as the required identifying assumptions are briefly outlined. To ease the burden of notation, the following shorthand notation will henceforth be used: $F_{Y_i \mid D_i=d, T_i=t}(y) := F_{Y_i \mid dt}(y)$, $F_{Y_i^0 \mid D_i=d, T_i=t}(y) := F_{Y_i^0 \mid dt}(y)$. Further, the τth quantile derived from an outcome distribution $F_{Y_i}(y)$ is interchangeably referred to as either $q_{\tau \mid dt}$ or $F^{-1}_{Y_i \mid dt}(\tau)$, with the potential quantile absent treatment arising analogously.

The QTT in the DiD context is generally defined as the discrete comparison of the factual and the counterfactually implied quantile for the treated in the post-treatment period

[5] Other approaches such as Callaway et al. (2018) and Callaway and Li (2019) require panel data and achieve identification by suitably restricting the joint distribution of the pooled distribution.

absent treatment:

$$QTT(\tau) = F^{-1}_{Y_i\,|\,11}(\tau) - F^{-1}_{Y_i^0\,|\,11}(\tau), \quad \tau \in (0,1). \tag{3.9}$$

By the same token, the DTT can be defined by comparing the factual with the potential outcome distribution along the support of the outcome variable $y \in \mathscr{Y} \subset \mathbb{R}$:

$$DTT(y) = F_{Y_i\,|\,11}(y) - F_{Y_i^0\,|\,11}(y). \tag{3.10}$$

In the following, DiD approaches that discuss the identifying assumptions – i.e., suitable common trend assumptions – for both (3.9) and (3.10) will be introduced.

The QDiD approach

Within any distributional DiD approach, the trend in potential outcomes does not only need to hold in mean outcomes but needs to hold more generally. Generally speaking, under a common trend assumption in quantiles, the counterfactual quantile for the treated in the post-treatment period can be expressed in terms of observable quantiles:

$$\begin{aligned} &F^{-1}_{Y_i^0\,|\,11}(\tau) - F^{-1}_{Y_i^0\,|\,10}(\tau) = F^{-1}_{Y_i^0\,|\,01}(\tau) - F^{-1}_{Y_i^0\,|\,00}(\tau) \\ &\implies F^{-1}_{Y_i^0\,|\,11}(\tau) = F^{-1}_{Y_i\,|\,10}(\tau) + \left(F^{-1}_{Y_i\,|\,01}(\tau) - F^{-1}_{Y_i\,|\,00}(\tau)\right), \quad \tau \in (0,1) \end{aligned} \tag{3.11}$$

Intuitively, what needs to be comparable for a *common trend in quantiles* to hold is that the trend in ranks must be the same across groups, while the level of the corresponding outcome is not restricted. Using a more structural view that involves the mapping from unobservables to the potential outcome space, Athey and Imbens (2002, 2006) show which conditions concerning the involved sub-populations are required for a common trend in quantiles to hold. As noted in Athey and Imbens (2002, p.16), the QDiD approach is only sensible if it is reasonable to assume that individuals are randomly assigned to a given group (d,t), i.e., one needs to face "identical populations of agents, subjected to different conditions in different groups and time periods".[6]

[6] Appendix B.1.1 provides a more formal in-depth consideration of the identifying assumptions using their structural setting.

From DTT to QTT: The DDiD approach

Another approach refers to Roth and Sant'Anna (2023) and Kim and Wooldridge (2024), which can be viewed as a more general distributional approach. There, the starting point is the identification of the entire counterfactual distribution of the treated population absent treatment, implying identification of the DTT. Thus, the approach can be referred to as a *distributional DiD (DDiD)* approach. In turn, this approach allows one to subsequently obtain distributional statistics, such as quantiles, from both the factual and the implied potential outcome distribution, allowing for the identification of the QTT. While Roth and Sant'Anna (2023) provide a general outline of the required assumptions for a distributional common trend to hold, Kim and Wooldridge (2024) explicitly consider the possibility to identify QTTs from counterfactual distributions within a DiD framework in which the distributional common trend can only be assumed to hold conditionally.

Akin to (3.11), a common trend assumption that holds along the entire support of the outcome variable, $y \in \mathscr{Y} \subset \mathbb{R}$, arises as

$$\begin{aligned} &F_{Y_i^0\,|\,11}(y) - F_{Y_i^0\,|\,10}(y) = F_{Y_i^0\,|\,01}(y) - F_{Y_i^0\,|\,00}(y) \\ &\implies F_{Y_i^0\,|\,11}(y) = F_{Y_i\,|\,10}(y) + \left(F_{Y_i\,|\,01}(y) - F_{Y_i\,|\,00}(y)\right). \end{aligned} \tag{3.12}$$

As formally shown in Roth and Sant'Anna (2023, Proposition 3.2), for a common trend to hold in population shares, it suffices to assume that the potential outcome for each group can be represented as follows:

$$F_{Y_i^0\,|\,dt}(y) = \theta \cdot G_t(y) + (1-\theta) \cdot H_d(y), \quad \text{for } y \in \mathscr{Y} \subset \mathbb{R},\ \theta \in [0,1], \text{ and } d,t \in \{0,1\}. \tag{3.13}$$

This implies that the common trend in population shares holds if the potential outcome for each group can be written as a combination of some fraction, $\theta \cdot G_t(y)$, for which treatment is essentially randomly assigned, and another fraction, $(1-\theta) \cdot H_d(y)$, being characterized by a stationary outcome.[7] That is, unlike in the QDiD case, it does not need

[7]The recent study by Kim and Wooldridge (2024) provides a milder interpretation under which a common trend in population shares holds. Particularly, they show that the condition requires that the distributional change (the "net flow of density", i.e., the time effect) is similar across groups absent treatment. This implies that for a common trend in population shares, it is *not* required that a fixed fraction of the population is characterized by stationarity/unconfoundedness.

to be the case that the sub-populations' agents are all identical and randomly subjected to a given (d,t) combination.

Once the counterfactual cdf in the post-treatment period is identified, taking the inverse of both the factual and the counterfactual cdf in (3.10) constitutes another way to derive the QTT.[8] One obvious advantage of this 'counterfactual cdf first' approach is the possibility to derive treatment effects for all sorts of distributional statistics beyond quantiles. This reasoning resembles the approaches put forward in, e.g., Rothe (2010) and Firpo and Pinto (2016), who both consider the identification of distributional effects of policies in the cross-sectional case. The DiD approach to identify QTTs proposed in Kim and Wooldridge (2024) is based upon the same rationale. It is crucial to note that the validity of the DDiD assumption does not imply the validity of the QDiD assumption. Intuitively, this is because of the non-applicability of the law of iterated expectations in the case of quantiles. This is shown more formally in appendix B.1.1.

3.4 RIF approaches and distributional DiD

As noted above, the RIF approach has recently found its way into the DiD literature. Below, to serve as a starting point for the more formal discussion, three exemplary recent articles are reviewed in greater detail against the backdrop of how RIF-DiD approaches are employed therein.

Evaluating the impact of a new sectoral minimum wage in Germany that affected only some occupational groups, Gregory and Zierahn (2022) used a RIF-DiD approach to examine the effect of this policy along the entire distribution of hourly wages. Doing so, they find very large effects especially for lower quantiles. In their main specification, they consider a 2×2 DiD framework in which they substituted the dependent variable with the RIF for various unconditional quantiles in their DiD regression formulation to identify the "unconditional quantile treatment effect (UQTE) of the minimum wage introduction" (Gregory and Zierahn, 2022, p.10). In describing their RIF approach, they note that they "define the influence function $IF(y;q^{\tau})$ of [their] wage outcome Y at sample quantile q^{τ}, which is then transformed (recentered) such that it aggregates back to the overall wage

[8] See Kim and Wooldridge (2024) for a more formal derivation.

distribution Y" (Gregory and Zierahn, 2022, p.9), which reads as if the RIF has been generated referring to all observations in their sample.

Another example of a RIF-DiD approach can be found in Bossler and Schank (2023). Within their broader characterization of drivers that underlie the more recent evolution of wage inequality in Germany, they put a particular focus on the impact of the newly introduced German minimum wage in 2015. For the latter, they resorted to a linear RIF-DiD regression formulation in which they additionally considered the dynamic effects of the minimum wage in the years prior to and after its introduction on various distributional statistics, including unconditional quantiles. To measure potential minimum wage effects, they make use of regional bite measures. Similar to Gregory and Zierahn (2022), their exposition of the RIF approach implies that they generate one RIF, such that their DiD interaction effect ultimately "identifies the average effect of the minimum wage bite at different locations of the unconditional distribution of log monthly wages after the minimum wage was introduced" (Bossler and Schank, 2023, p.16).[9]

In another contribution, Havnes and Mogstad (2015) examine the effect of the introduction of universal child care programs in Norway on the later earnings distribution. Even though they formally resemble Gregory and Zierahn (2022) – i.e., employing a 2×2 DiD regression formulation – their outline scrutinizes the RIF mechanics to a larger extent. While Gregory and Zierahn (2022) and Bossler and Schank (2023) explicitly framed the RIF-DiD approach as one way of directly estimating unconditional quantile treatment effects, Havnes and Mogstad (2015) pursue a somewhat different approach in that they first define how the DiD notion leads to a counterfactual distribution that can be used to derive quantile effects in a second step. Concretely, they start by showing that under a distributional common trend assumption – similar to Roth and Sant'Anna (2023) and Kim and Wooldridge (2024) – a counterfactual distribution in the post-treatment period absent treatment can be identified. This distributional DiD approach ultimately boils down to first estimating an effect on population shares and then to "divide [this population share

[9]For the sake of completeness, note that three more recent contributions (Caliendo et al., 2023; Bossler et al., 2024; Pusch, 2024) also combine a RIF-regression with a DiD approach using bite measures to identify the distributional treatment effect of the minimum wage introduction in Germany. They do not provide any detailed formal outline of their approach other than that they make use of a linear DiD regression specification with the RIF of their outcome variable as the dependent variable. Therefore, as in Gregory and Zierahn (2022) and Bossler and Schank (2023), these approaches most likely derive a single RIF from the pooled distribution encompassing all the years considered.

effect] (...) by a kernel estimate of the joint density of earnings (...) to arrive at the associated QTE." (Havnes and Mogstad, 2015, p. 106).

In the following, it is shown that the formal RIF-DiD outlines of these contributions, in parts, neglect certain obstacles and shortcomings of a RIF-DiD approach. Briefly summarized, the first two approaches follow a 'pooled' RIF-DiD approach that – as will be described below – results in misleading approximations of quantile treatment effects. On the other hand, the account by Havnes and Mogstad (2015) provides the groundwork for the 'correct' RIF-DiD approach that is proposed below. However, their approach makes use of a slightly different kernel density to estimate the quantile treatment effect, as the discussion below will highlight.

3.4.1 The pooled RIF-DiD approach and corresponding pitfalls

Based on the outlined examples, why would employing a pooled RIF approach be problematic? To see why, recall that in the simplest 2×2 DiD case, the non-parametric representation of the average treatment effect on the treated can be easily obtained using regression techniques by considering a saturated linear model. This familiar DiD regression formulation is given by

$$\mathbb{E}[Y \,|\, D,T] = \alpha + \beta\, D + \delta\, T + \gamma\,(D \times T). \tag{3.14}$$

The fact that a RIF procedure allows for a representation of, e.g., unconditional quantiles in terms of an expected value as in (3.5) may, at first sight, suggest that it is possible to identify the treatment effect on the treated for other distributional statistics such as unconditional quantiles by simply using the corresponding RIF instead of Y in (3.14). Concretely, this rationale involves generating the RIF for the τth quantile according to (3.4), and subsequently estimating the treatment coefficient by OLS using the RIF as the dependent variable. In such an approach, $\widehat{\gamma}$ would be interpreted as an estimator for the QTT.

In what follows, it is demonstrated why such an approach is misleading due to the aforementioned issue of a left-hand side misspecification. To illustrate this, replace Y with

$RIF(y;q_\tau,F_Y)$, and note that the RIF in this particular case reads

$$RIF(y;q_\tau^m,F_{Y^m}) = q_\tau^m + \frac{\tau - \mathbb{1}[y \leq q_\tau^m]}{f_{Y^m}(q_\tau^m)}, \tag{3.15}$$

where F_{Y^m} refers to the cdf and f_{Y^m} refers to the pdf of the *pooled* distribution comprising both the pre- and post-treatment period as well as treated and non-treated individuals. Additionally, q_τ^m refers to the unconditional τth quantile from the *pooled* distribution. The conditional expected value of the RIF in (3.15) for group $(D,T)=(d,t)$ is given by:

$$\mathbb{E}[RIF(Y;q_\tau^m,F_{Y^m})\,|\,D=d,T=t] = q_\tau^m + \frac{\tau - \mathbb{P}[Y \leq q_\tau^m\,|\,D=d,T=t]}{f_{Y^m}(q_\tau^m)}.$$

Considering a linear DiD specification, a pooled RIF-DiD regression would therefore read

$$q_\tau^m + \frac{\tau - \mathbb{P}[Y \leq q_\tau^m\,|\,D=d,T=t]}{f_{Y^m}(q_\tau^m)} = \alpha_\tau^{RIF} + \beta_\tau^{RIF}\,D + \delta_\tau^{RIF}\,T + \gamma_\tau^{RIF}\,(D\times T),$$

involving a distribution regression idea in the numerator, i.e.,

$$\mathbb{P}[Y \leq q_\tau^m\,|\,D,T] = \alpha_\tau^{DR,m} + \beta_\tau^{DR,m}\,D + \delta_\tau^{DR,m}\,T + \gamma_\tau^{DR,m}\,(D\times T).$$

Crucially, as noted in (3.8), the RIF coefficient of interest, γ_τ^{RIF}, contains the 'horizontal translation' idea formulated above, i.e., the RIF regression coefficients can be directly linked to the DR coefficients:

$$\gamma_\tau^{RIF} = -\frac{\gamma_\tau^{DR,m}}{f_{Y^m}(q_\tau^m)}. \tag{3.16}$$

It needs to be noted that the involved terms all concern a quantile that is derived from the pooled distribution. Thus, the implied effect that could be estimated by γ_τ^{RIF} turns out to be *technically* misleading for an estimation of the QTT. Formally, the involved counterfactual quantile that could be recovered using a pooled RIF-DiD approach arises as follows:

$$\begin{aligned}\mathbb{E}[RIF(Y;q_\tau^m,F_{Y^m}(y))\,|\,D=1,T=1] &= q_\tau^m + \frac{\tau-\mathbb{P}[Y\leq q_\tau^m\,|\,D=1,T=1]}{f_{Y^m}(q_\tau^m)}\\ &= \underbrace{q_\tau^m + \frac{\tau-(\alpha_\tau^{DR,m}+\beta_\tau^{DR,m}+\delta_\tau^{DR,m})}{f_{Y^m}(q_\tau^m)}}_{\hat{=}\widetilde{q}_\tau^{0,m}} - \frac{\gamma_\tau^{DR,m}}{f_{Y^m}(q_\tau^m)},\end{aligned}$$

where $\widetilde{q}_\tau^{0,m}$ denotes the RIF-based approximation of the counterfactual quantile. The *supposed* QTT that could be approximated hence actually coincides with

$$\widetilde{QTT}_\tau^{pooled} = -\frac{\gamma_\tau^{DR,m}}{f_{Y^m}(q_\tau^m)} = \underbrace{q_\tau^m + \frac{\tau-\mathbb{P}[Y\leq q_\tau^m\,|\,D=1,T=1]}{f_{Y^m}(q_\tau^m)}}_{\substack{=\mathbb{E}[RIF(Y;q_\tau^m,F_{Y^m}(y))\,|\,D=1,T=1]\\ \text{(i)}}} - \underbrace{\widetilde{q}_\tau^{0,m}}_{\text{(ii)}}. \tag{3.17}$$

To disentangle what the pooled RIF-DiD of (3.17) is ultimately able to approximate, it is useful to scrutinize the involved terms on which the supposed treatment effect approximation is based on. Note first that a credible estimate/approximation of the QTT requires that the implied terms of the pooled RIF approach coincide, at least approximately, with correct terms that are involved in the QTT from (3.11). Looking at the given terms in (3.17), this is certainly not the case: The first term, (i), is some sort of RIF transformation around the τth quantile of the *pooled* distribution.[10] Crucially, this term does not refer to a valid expression for $F^{-1}_{Y\,|\,11}$ from (3.9). That is, this RIF expression does *not* return the quantile from the $(D,T)=(1,1)$ group as the RIF is not defined for this quantile. Moreover, the second term, (ii), is not a valid representation of the required counterfactual quantile derived from the $(D,T)=(1,1)$ group since it also revolves around the pooled quantile, not $q_{\tau\,|\,11}$. Beside the fact that the approximation of the actual effect might be poor, the point around which one approximates the apparent QTT is not meaningful in the first place, somewhat obviating a detailed discussion of the required identifying assumptions. Overall, the supposed treatment effect estimate involves some peculiar RIF transformation that does not contain any reasonable statements that relate to the QTT as specified in (3.9).

[10]Technically, it is the RIF transformation of the deviation of τ from the implied probability of falling below the pooled quantile for individuals in group $(D,T)=(1,1)$, using the slope around the τth quantile from the pooled distribution.

3.4.2 Overcoming the pooled RIF-DiD shortcomings: DDiD- and QDiD-RIF

The formulation in (3.16) makes the issues with a pooled approach obvious: The approximation of the supposed quantile treatment effect is structurally misleading since the effect is obtained using quantiles that refer to the pooled distribution. In the following, two approaches are proposed that overcome this shortcoming.

The "correct" RIF-DiD approximation: DDiD-RIF

The first approach is based on the RIF mechanism of translating vertical (population share) into horizontal (quantile) shifts described above (see figure 3.1). This reasoning can be used to approximate effects around the *correct* quantile of interest, i.e., $q_{\tau\,|\,11}$. As will be shown below, this 'correct' RIF-DiD approximation idea relies on the DDiD assumptions outlined above, and thus this approach will henceforth be referred to as DDiD-RIF.

To see how the QTT can be approximated using a RIF approach, consider the following conditional expected value of an adjusted RIF, which approximates effects around $q_{\tau\,|\,11}$:

$$\begin{aligned}\mathbb{E}[RIF(Y;q_{\tau\,|\,11},F_{Y\,|\,11}(y))\,|\,D=1,T=1] &= q_{\tau\,|\,11}+\frac{\tau-\mathbb{P}[Y\leq q_{\tau\,|\,11}\,|\,D=1,T=1]}{f_{Y\,|\,11}(q_{\tau\,|\,11})}\\ &= \underbrace{q_{\tau\,|\,11}+\frac{\tau-(\alpha_\tau^{DR}+\beta_\tau^{DR}+\delta_\tau^{DR})}{f_{Y\,|\,11}(q_{\tau\,|\,11})}}_{\equiv\widetilde{q}^{0}_{\tau\,|\,11}}-\frac{\gamma_\tau^{DR}}{f_{Y\,|\,11}(q_{\tau\,|\,11})},\end{aligned} \tag{3.18}$$

with $\widetilde{q}^{0}_{\tau\,|\,11}$ being the DDiD-RIF-based approximation of the τth counterfactual quantile of the treated in the post-treatment period. Following the QTT definition from (3.9), it is apparent that the RIF approximation now indeed entails a valid formulation for a QTT approximation. Rewriting (3.18) by bringing $\tilde{q}^0_{\tau\,|\,11}$ to the other side yields:[11]

$$-\frac{\gamma_\tau^{DR}}{f_{Y\,|\,11}(q_{\tau\,|\,11})} = q_{\tau\,|\,11} - \tilde{q}^0_{\tau\,|\,11} = \widetilde{QTT}_\tau^{DDiD}. \tag{3.19}$$

The involved rationale is reminiscent of approaches that first identify the entire counterfactual cdf to derive quantiles from the factual and the counterfactual distribution to get an estimate for the QTT. The DDiD-RIF approach is based on the same notion but utilizes first-order approximations, i.e., local inversions, instead of globally inverting the counterfactually-implied cdf as proposed by Biewen et al. (2022) or Kim and Wooldridge (2024).[12] The following reformulation of (3.18) makes this obvious:

$$-\frac{\gamma_\tau^{DR}}{f_{Y\,|\,11}(q_{\tau\,|\,11})} = -\frac{\overbrace{\mathbb{P}[Y \leq q_{\tau\,|\,11}\,|\,D=1,T=1]}^{=\tau} - \overbrace{\mathbb{P}[Y^0 \leq q_{\tau\,|\,11}\,|\,D=1,T=1]}^{=(\alpha^{DR}+\beta^{DR}+\delta^{DR})}}{f_{Y\,|\,11}(q_{\tau\,|\,11})},$$

implying that the DDiD-RIF incorporates the identification of a counterfactual cdf of the treated absent treatment in the post-treatment period evaluated at $q_{\tau\,|\,11}$. The $\Delta\tau$ in this case entails the comparison between the factual (τ) and the model-implied counterfactual cdf ($\alpha^{DR}+\beta^{DR}+\delta^{DR}$). This population share difference is then translated into a quantile effect, as illustrated in figure 3.2.

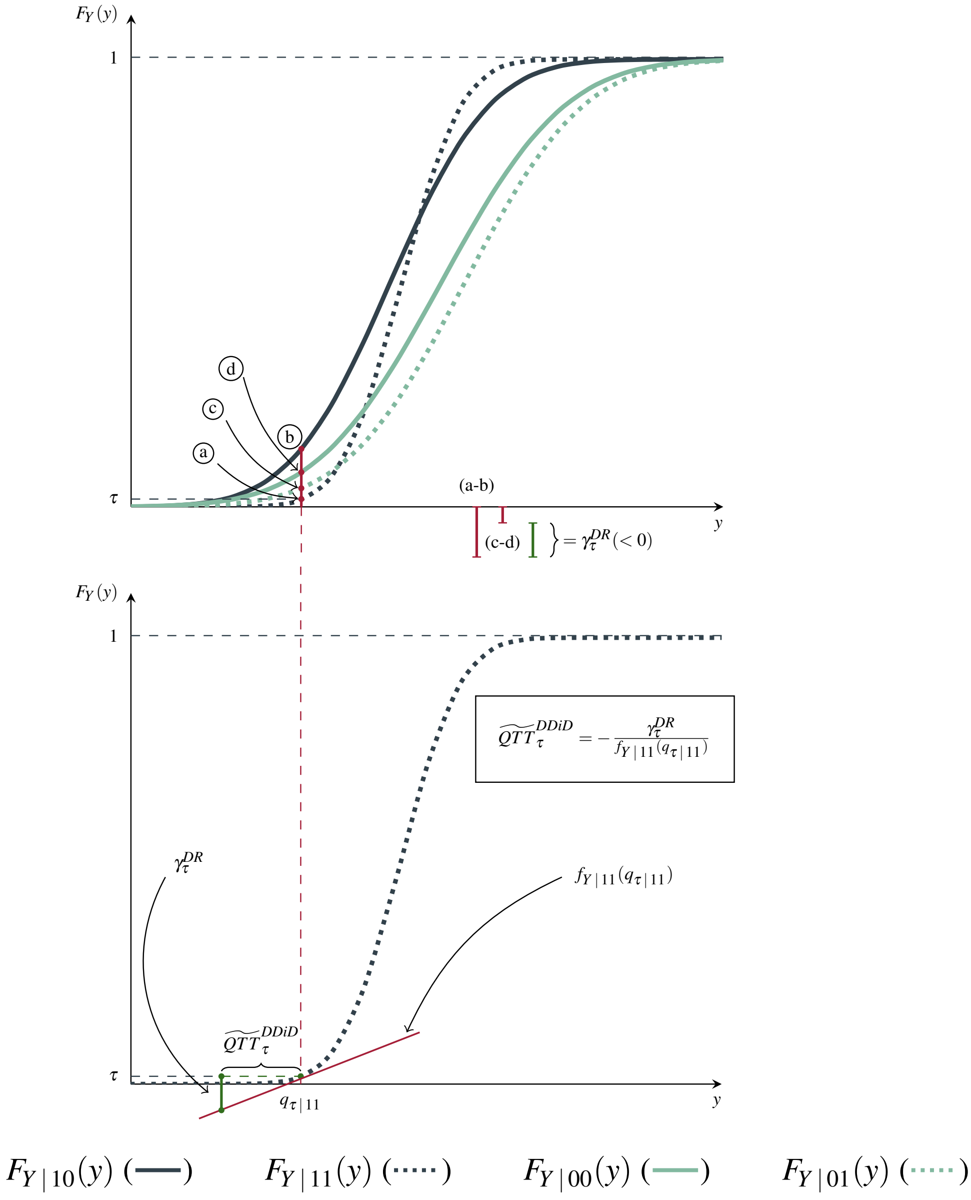
To summarize, it is evident that the DDiD-RIF approach does not require a common trend assumption in quantiles. Instead, it relies on the DDiD assumptions to identify the counterfactual population share in the numerator. The translation of the population share treatment effect into a quantile effect can be viewed as a 'mechanical' step that depends not on structural identifying assumptions, but on technical ones such as smooth outcome distributions and the absence of mass points (see discussion above). One remaining short-

[11] Note that in a saturated model as considered here, it holds that $\mathbb{P}[Y \leq q_{\tau\,|\,11}\,|\,D=1,T=1] = \tau$ since the probability around the τth quantile in $(D,T) = (1,1)$ is considered – not the quantile from the pooled distribution. Thus $q_{\tau\,|\,11} + \left(\tau - \mathbb{P}[Y \leq q_{\tau\,|\,11}\,|\,D=1,T=1]\right)/f_{Y\,|\,11}(q_{\tau\,|\,11}) = q_{\tau\,|\,11}$ since the numerator of the second term becomes zero.

[12] Kim and Wooldridge (2024) develop a theoretical framework to estimate QTTs by means of inverting factual and counterfactual cumulative distribution functions, with the latter being identified by a common trend in population shares (see also appendix B.2). Biewen et al. (2022) study the distributional effects of a minimum wage introduction. Their methodological approach involves a full model of the counterfactual cdf and the subsequent global inversion thereof to estimate, inter alia, QTTs in a DiD framework. As in Kim and Wooldridge (2024), they rely on a common trend in population shares to identify the distributional treatment effects.

coming of this approach is that the true QTT may be poorly approximated, which is an inherent disadvantage of RIF-based methods.

Figure 3.2: DDiD-RIF translation – graphical intuition



Notes: Figure shows the translation mechanism on which the DDiD-RIF approach as formulated in (3.19) is based. The upper figure indicates the 'vertical' effect on population shares implied by the treatment in the post-treatment period evaluated at the τth quantile of the group $(D,T) = (1,1)$. This vertical effect is captured by the distribution regression coefficient γ_τ^{DR}. The translation on the horizontal axis is operationalized by virtue of dividing the vertical distance using the tangent line evaluated at $q_{\tau|11}$, i.e., $f_{Y|11}(q_{\tau|11})$. As outlined in figure 3.1, the translation is an application of the triangular formula. However, now it is not a matter of translating a covariate-induced shift in τ, but rather the shift associated with the treatment in the post-treatment period.

Finally, to operationalize this approach, one can follow a simple two-step procedure, using easily implementable distribution regression techniques (Chernozhukov et al., 2013):

(i) Estimate the treatment effect on the population share falling below the quantile of interest $q_{\tau\,|\,11}$, using the following estimation equation:

$$\mathbb{1}[Y_i \leq q_{\tau\,|\,11}] = \alpha_\tau^{DR} + \beta_\tau^{DR}\, D_i + \delta_\tau^{DR}\, T_i + \gamma_\tau^{DR}\,(D_i \times T_i) + \varepsilon_i.$$

Estimates can, e.g., be obtained by means of OLS.[13]

(ii) Obtain a QTT approximation by dividing the estimate $\widehat{\gamma}_\tau^{DR}$ by an estimate of the density around the quantile of interest, $\widehat{f}_Y(\widehat{q}_{\tau\,|\,11})$:[14]

$$\widehat{\widehat{QTT}}_\tau^{DDiD} = -\frac{\widehat{\gamma}_\tau^{DR}}{\widehat{f}_{Y\,|\,11}(\widehat{q}_{\tau\,|\,11})}. \tag{3.20}$$

As noted above, a similar approach has been used in Havnes and Mogstad (2015), however, instead of using the density estimator of the sub-population $(D,T) = (1,1)$, they make use of the "kernel estimate of the joint density of earnings" (Havnes and Mogstad, 2015, p.106). Their approach has been criticized by Kottelenberg and Lehrer (2017), arguing that the RIF approach is developed for cross-sectional settings and thus does not easily extend to the case involving several time periods, which is reminiscent to the critique formulated above. Kottelenberg and Lehrer (2017) further claim that the required assumptions to point identify the QTTs underlying this approach are the ones formulated in Athey and Imbens (2006), i.e., identical population of agents subjected to different conditions in different groups and over time. While this is true for the case in which one actually compares the quantiles with one another to leverage a common trend in quantiles (see QDiD-RIF case below), this stands in contrast to the established result that such an

[13]Note that, in applications in which the common trend can only be assumed to hold conditional on some covariates, one could – beside using additive controls in the DiD regression formulation – also employ a reweighting approach as in Kim and Wooldridge (2024). This involves the computation of weights, $\widehat{\omega}_j(x)$, in a first stage and estimate the parameters from the linear DiD regression model using WLS. Appendix B.2 provides an outline of their approach and how one could make use of the implied weights in a RIF-DiD approach.

[14]The quantile estimate $\widehat{q}_{\tau\,|\,11}$ is obtained by considering only the relevant sub-population, i.e., $\widehat{q}_{\tau\,|\,11} = \arg\min_u \sum_{i\in(11)}(\tau - \mathbb{1}[Y - u \leq 0]) \cdot (Y - u)$.

approach actually relies on the DDiD assumptions outlined above.[15]

An additional advantage of the DDiD-RIF approach is that this principle can also be used to not only identify the QTT but also to identify the quantile treatment effect for the entire sub-population in $T = 1$, i.e., the QTE_1. Intuitively, this is feasible since the identification part concerns expected values of indicator variables, allowing for the application of the law of total expectations. The fundamental idea for a slightly altered DDiD-RIF approach to identify the QTE_1 is outlined in appendix B.1.2.

The QDiD-RIF approach

The second proposed RIF-DiD approach revolves around the fact that the RIF is a way to express the unconditional quantile in terms of an expected value, as in (3.2). As the discussion below will highlight, this approach relies on a quantile common trend assumption. Therefore, it will henceforth be referred to as QDiD-RIF.

The QDiD-RIF is based on a compound RIF used as the new dependent variable in a linear DiD regression formulation, each capturing the respective RIF of the involved sub-populations. This encompasses the idea of updating the RIF to the corresponding sub-populations, which also overcomes the issue of approximating around misleading quantiles.[16] Hence, one could view this as constituting a bridge between a classical RIF and an inequality treatment effect (ITE, Firpo and Pinto, 2016) approach, as outlined in Rios-Avila and Maroto (2024) for the cross-sectional case. The corresponding DiD formulation then reads

$$\begin{aligned}
&\mathbb{E}[RIF_\tau^{comp} \,|\, D,T] = \alpha_\tau + \beta_\tau D + \delta_\tau T + \gamma_\tau (D \cdot T), \quad \text{with} \qquad (3.21)\\
&RIF_\tau^{comp} = (D \cdot T) \times RIF(y; q_{\tau\,|\,11}, F_{Y\,|\,11}(y)) + (D \cdot (1-T)) \times RIF(y; q_{\tau\,|\,10}, F_{Y\,|\,10}(y))\\
&\quad + ((1-D) \cdot T) \times RIF(y; q_{\tau\,|\,01}, F_{Y\,|\,01}(y)) + ((1-D) \cdot (1-T)) \times RIF(y; q_{\tau\,|\,00}, F_{Y\,|\,00}(y)).
\end{aligned}$$

Since (3.21) is saturated, γ_τ refers to a double difference incorporating the discrete com-

[15] Indeed, Havnes and Mogstad (2015) argue as well that they need to invoke a common trend in population shares rather than in quantiles.

[16] The notion of the necessity of 'updating' the RIF to avoid the approximation error being too stark has also been made in Rios-Avila and de New (2022).

parison of sub-population-specific quantiles. To see this, note:

$$
\begin{aligned}
\mathbb{E}[RIF_{\tau}^{comp} \,|\, D=1,T=1] &= \mathbb{E}[RIF(Y;q_{\tau\,|\,11},F_{Y\,|\,11}(y))\,|\,D=1,T=1] \\
&= \alpha_\tau+\beta_\tau+\delta_\tau+\gamma_\tau, \\
\mathbb{E}[RIF_{\tau}^{comp} \,|\, D=1,T=0] &= \mathbb{E}[RIF(Y;q_{\tau\,|\,10},F_{Y\,|\,10}(y))\,|\,D=1,T=0] = \alpha_\tau+\beta_\tau, \\
\mathbb{E}[RIF_{\tau}^{comp} \,|\, D=0,T=1] &= \mathbb{E}[RIF(Y;q_{\tau\,|\,01},F_{Y\,|\,01}(y))\,|\,D=0,T=1] = \alpha_\tau+\delta_\tau, \\
\mathbb{E}[RIF_{\tau}^{comp} \,|\, D=0,T=0] &= \mathbb{E}[RIF(Y;q_{\tau\,|\,00},F_{Y\,|\,00}(y))\,|\,D=0,T=0] = \alpha_\tau,
\end{aligned}
$$

with $\mathbb{E}[RIF(Y;q_{\tau\,|\,dt},F_{Y\,|\,dt}(y))\,|\,D=d,T=t] = q_{\tau\,|\,dt}$ by virtue of (3.2), and hence

$$
\widetilde{QTT}_{\tau}^{QDiD} = \gamma_\tau = \left(q_{\tau\,|\,11}-q_{\tau\,|\,10}\right)-\left(q_{\tau\,|\,01}-q_{\tau\,|\,00}\right). \tag{3.22}
$$

As in the DDiD-RIF case, an estimate for γ_τ in (3.21) can be obtained using OLS. Taken together, it is apparent that for γ_τ to correspond to the QTT, a common trend in quantiles, as defined in (3.11), is required. Figure 3.3 illustrates the involved rationale graphically.[17]

In fact, in the absence of covariates, the outlined RIF-QDiD approach in the simple 2×2 case leads to results equivalent to those of a CQR approach (Koenker and Bassett Jr., 1978), since both approaches represent a discrete comparison of sub-population-specific quantiles. Note, however, as mentioned in Dube (2019b), as soon as one would like to control for covariates in the linear regression formulation, the two approaches differ: While the RIF-QDiD approach allows to *control* for covariates, leaving the interpretation of the quantiles as *unconditional* quantiles untouched, the CQR approach alters the interpretation of the treatment effect such that the treatment effect now is contingent on the included covariates.[18] As a matter of fact, the QDiD-RIF approach has already been briefly mentioned as an ad hoc solution by Rios-Avila (2020) but was, however, not clearly related to the underlying identifying assumptions.

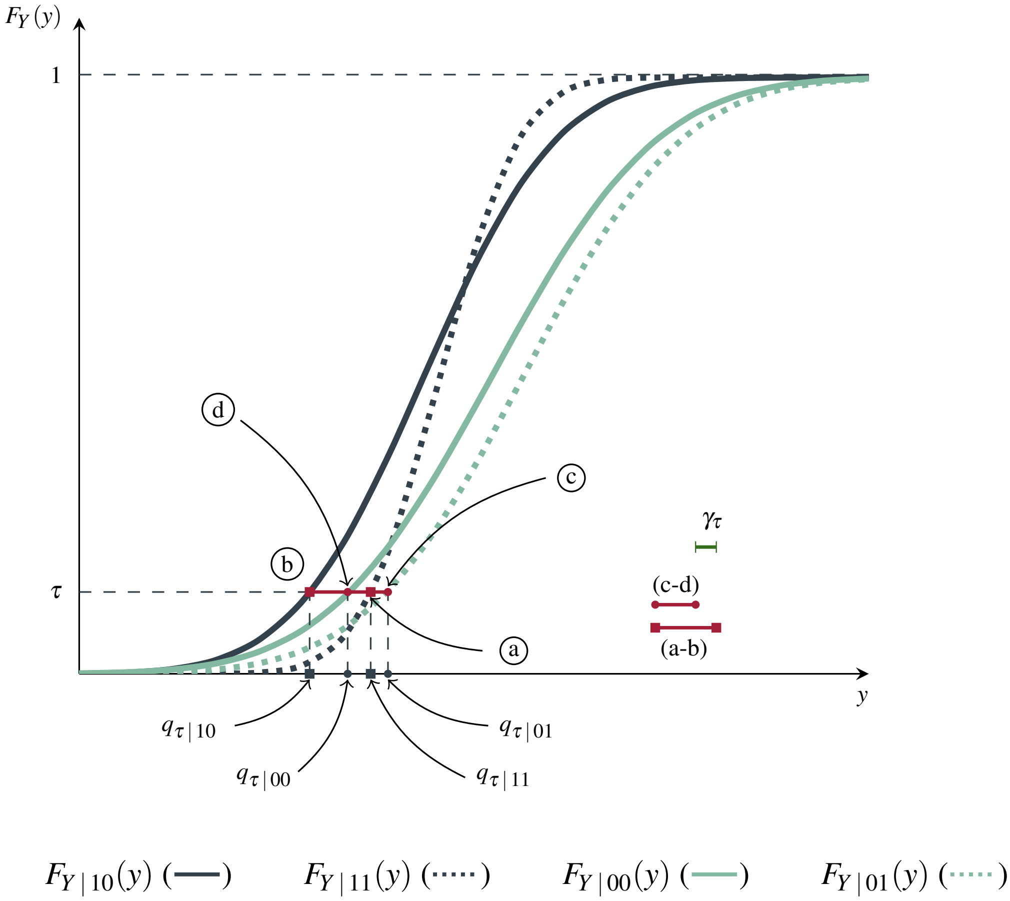
Note that since this approach is based on a quantile common trend assumption, it is not possible to obtain QTE_1 via the QDiD-RIF approach, as was possible in the DDiD-RIF

[17]See, e.g., Athey and Imbens (2006) and Havnes and Mogstad (2015) for comparable graphical illustrations of the QDiD rationale.

[18]Beside the ability to linearly control for confounding covariates by means of including them in a linear regression as additional controls, one could also follow the rationale put forward in Kim and Wooldridge (2024) and reweight each sub-population accordingly using the reweighting factors $\widehat{\omega}_j(x)$ as described in appendix B.2 to adjust for covariates.

case, since the law of total expectations does not apply to quantiles. Furthermore, estimating QTTs in settings with multi-level treatment and/or multiple time periods is not always feasible, as RIFs for each implied sub-population would need to be formed.

Figure 3.3: QDiD-RIF – graphical intuition



Notes: The figure shows how the QTT is recovered when employing the QDiD-RIF approach. As illustrated, the QDiD-RIF approach entails a discrete comparison of sub-population-specific quantiles to leverage a common trend in the τth quantile. This discrete comparison of quantiles is incorporated in γ_τ which comprises the double difference in quantiles as indicated in (3.22).

3.5 Monte Carlo simulation study

In the following, the results of a Monte Carlo (MC) simulation study are examined to compare the performance of the methods discussed above. To focus on shortcomings that are inherent to the respective RIF approach, no covariates are considered in the data generating process (DGP). Moreover, before considering the actual simulation results, sources of potential biases arising from mere formal considerations are briefly described to assess the simulation results against this background.

Abstracting from the described RIF-inherent shortcomings, such as potentially substantial approximation errors, the primary driver of bias in DiD frameworks boils down to a given approach's ability to accurately model a counterfactual outcome absent treatment. To illustrate this, the bias of the non-parametric double-difference DiD estimator for some statistic, $\varphi(\cdot)$, is examined in greater detail. The non-parametric DiD estimator for $\varphi(\cdot)$ reads

$$\widehat{\psi}_{DiD}(\varphi) = [\widehat{\varphi}(Y_{11}) - \widehat{\varphi}(Y_{10})] - [\widehat{\varphi}(Y_{01}) - \widehat{\varphi}(Y_{00})], \tag{3.23}$$

with $\widehat{\varphi}(\cdot)$ being an unbiased estimator for $\varphi(\cdot)$, and $\varphi(\cdot)$ being the statistic for which a common trend in potential outcomes is assumed, e.g., the mean, quantiles, or the population share. The true, oracle, treatment effect is given by

$$\psi_0(\varphi) = \varphi(Y_{11}) - \varphi(Y_{11}^0). \tag{3.24}$$

Therefore, the bias of this DiD estimator arises as

$$\mathbb{E}[\widehat{\psi}_{DiD}(\varphi) - \psi_0(\varphi)] = \left[\varphi(Y_{11}^0) - \varphi(Y_{10})\right] - [\varphi(Y_{01}) - \varphi(Y_{00})]. \tag{3.25}$$

This formalizes the insight that any approach failing to accommodate the required common trend assumption will inevitably lead to biased results. Another, more subtle insight is that the bias increases with the magnitude of the common trend violation. Put differently, the larger the deviation in the temporal differences of potential outcomes, the greater the bias will be if the common trend assumption is not met. On the other hand, the bias of the DiD estimator is immune to variations in the intensity of the treatment effect.

The data generating process described below is constructed in such a way that it accommodates the DDiD, but not the QDiD assumptions that were outlined previously. Following the formal reasoning above, one would thus expect the QDiD-RIF approach to produce biased QTT estimates under the given DGP, increasingly so with higher common trend intensities, while no effect on the bias from higher treatment intensities is anticipated. Notably, approximation errors are irrelevant for the QDiD-RIF approach, as the

RIFs are 'updated'. Conversely, regarding the bias of results produced by the DDiD-RIF approach – which adheres to the underlying common trend assumption implied by the DGP – it is expected that any bias solely arises from the inherent approximative issues of the RIF. Put differently, the only source of bias in the DDiD-RIF approach is the treatment effect intensity, while this approach remains unaffected by large common trend deviations. Finally, the discussion above has highlighted that the bias associated with estimates obtained using the pooled RIF approach needs to be rationalized from a somewhat 'mechanical' perspective: The greater the difference between the pooled outcome distribution and the marginal $(D,T) = (1,1)$ subpopulation distribution, the greater the bias. This can be ascribed to the discrepancy between the quantiles around which a pooled RIF approach approximates the effects and the actual quantiles of interest for the QTT. In turn, this renders the pooled RIF's bias susceptible to both common trend and treatment intensity.[19]

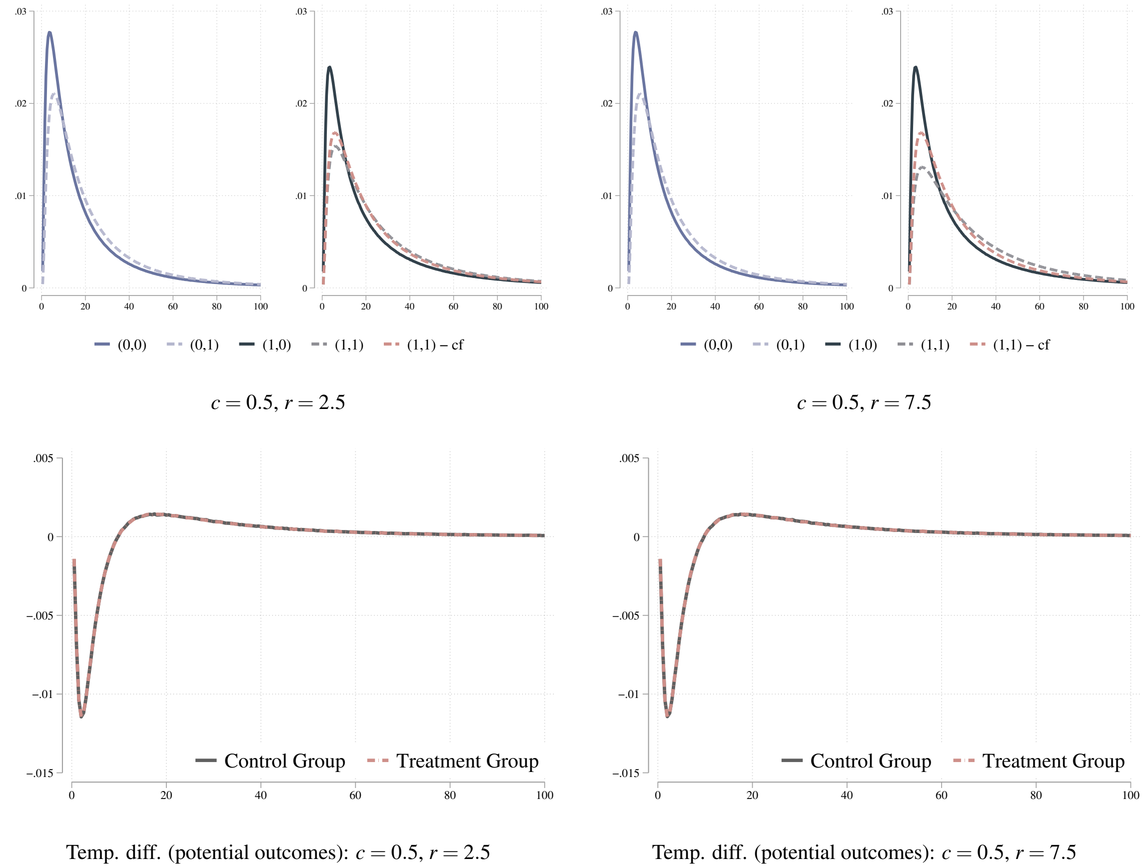
To be able to carve out the extent of the bias in cases in which a common trend in population shares but not in quantiles is supported, the aforementioned mixture distribution given in (3.13) and described in Roth and Sant'Anna (2023, Example 3, p.741) is used to formulate the DGP. In contrast to Roth and Sant'Anna (2023), the potential outcome distribution for individuals in group $D_i = d$ at time $T_i = t$ in this study additionally incorporates varying common trend intensities:

$$\begin{aligned} F_{Y_i^0|dt} &= \theta \times G_{it} + (1-\theta) \times H_{id}, \qquad (3.26) \\ \text{with} \quad G_{it} &\sim \text{lognormal}(2 + c \cdot \mathbb{1}[T_i = t], 1), \\ \text{and} \quad H_{id} &\sim \text{lognormal}(3 + c \cdot \mathbb{1}[D_i = d], 1), \end{aligned}$$

where $c \in \{0.1, 0.25, 0.5, 0.75\}$ refers to the extent to which potential distributions change in the absence of treatment. Below, the balanced case with $\theta = 0.5$ is considered. As to the sample size, 100,000 observations for each of the four groups are sampled according to (3.26), resulting in an equal share of control/treated as well as pre-/post-period observations. This setup creates a difference in levels among the four involved sub-populations, but it ensures that the potential outcome distributions for both the treatment and control groups follow the same trend over time in terms of potential population shares.

[19] Also compare the discussion of the RIF employed by Dube (2019b) regarding the appropriateness of the RIF approach applied therein when facing multi-period data.

Figure 3.4: Distributions of actual and potential outcomes/temporal differences in distributional mass of potential outcomes



Notes: The figure's first row displays pdfs for all four sub-populations, including the potential distribution absent treatment for the $(D,T) = (1,1)$ sub-population. Pdfs are computed using a bin-wise consideration, i.e., the distributional mass for a given threshold, z_j, is given by $\mathbb{P}[Y \leq z_j] - \mathbb{P}[Y \leq z_{j-1}]$, where the increments between thresholds are 0.5. The second row shows the temporal differences in distributional masses of potential outcomes for both the treated and the control group with $N = 100,000$ per sub-population, sampled 500 times. Pdfs and differences thereof hence refer to the average over all 500 MC samples. Figures are based on the DGP as specified in (3.26) and (3.27) with an exemplifying common trend intensity of $c = 0.5$ and a treatment intensity of $r = 2.5$ for the first column and $r = 7.5$ for the second column, respectively. Considering the second row, the figure shows that the specified DGP indeed satisfies a common trend in population shares (cf., Roth and Sant'Anna, 2023, Figure 1 therein).

Treatment is modeled as rank-dependent and non-linear, implying a varying treatment effect across different regions of the distribution and, crucially, not merely a location shift but also a change in the shape of the resulting outcome distribution.[20] The outcome for the treated group in the treatment period, $Y_i(1,1)$, is given by

$$Y_i(1,1) = Y_i^0(1,1) + \left(r \times \left[R_i + (1/2)\cdot R_i^2 + (1/3)\cdot R_i^3\right]\right), \tag{3.27}$$

with $R_i \in [0,1]$ describing the rank of individual i in the potential outcome distribution absent treatment for the treated in the post-treatment period, $F_{Y_i^0|11}$. For example, individuals around the 50th percentile of the potential outcome distribution get assigned a value of $R_i \approx 0.5$. The distribution of both actual and potential outcomes, as well as the temporal differences in potential outcomes, are illustrated in figure 3.4, highlighting that the DGP implies a common trend in population shares.

This setup implies that the oracle treatment effect for the τth quantile can be computed exactly as it only depends on the known parameters τ (the rank in the population) and the treatment intensity r.[21] Similar to the common trend intensity, the treatment intensity is governed by means of the scalar $r \in \{2.5, 5, 7.5\}$. Given this flexible specification involving various common trend and treatment intensities, it is possible to examine the considered RIF methods regarding their sensitivity to smaller or larger differences over time or across group, making it possible to validate the ex ante assertions from above.

Simulation results: Implications of the DGP

To generally appraise the implications of the DGP specified in (3.26) and (3.27) for estimates of distributional treatment effects, table 3.1 presents baseline findings on non-parametric DiD estimates for both quantiles and population share treatment effects. Furthermore, table 3.1 provides a summary of sub-population-specific quantiles and population shares, group-specific time trends in potential outcomes, and the oracle treatment effects on the treated. Using the oracle treatment effect, the table also displays the bias and the RMSE of the non-parametric DiD estimator for quantile and population share effects (see table notes for further technical details).

[20] The idea for the rank dependence of treatment effects is taken from Borgen et al. (2022).

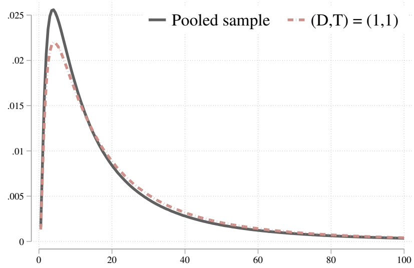
[21] To see this, rewrite (3.27) for the τth quantile of the outcome distribution Y_i, with R_i in this case being equal to τ (corresponding to a population perspective): $QTT_\tau \equiv q_{\tau|11} - q^0_{\tau|11} = r \times \left[\tau + (1/2)\cdot\tau^2 + (1/3)\cdot\tau^3\right]$, $\tau \in [0,1]$.

Table 3.1: Implications of the DGP for distributional DiD estimates

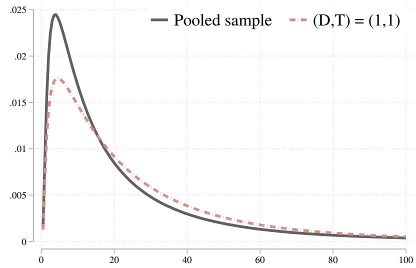
	Y_{11}^0	Y_{11}	Y_{10}	Y_{01}	Y_{00}	$\Delta_t Y_{D=1}^0$	$\Delta_t Y_{D=0}$	ψ_0	$\widehat{\psi}_{DiD}$	Bias	RMSE
c = 0.1, r = 2.5											
q_{10}	3.192	3.456	2.941	3.129	2.888	0.251	0.241	0.263	0.273	0.010	0.037
q_{50}	13.459	15.126	12.808	12.809	12.181	0.651	0.628	1.667	1.690	0.023	0.114
q_{90}	56.775	60.645	55.771	52.450	51.370	1.004	1.080	3.870	3.794	-0.076	0.646
$F_{Y\mid DT}(q_{10\mid 11})$	0.113	0.100	0.127	0.117	0.131	-0.014	-0.014	-0.013	-0.013	0.000	0.002
$F_{Y\mid DT}(q_{50\mid 11})$	0.541	0.500	0.557	0.560	0.576	-0.016	-0.016	-0.041	-0.041	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.910	0.900	0.913	0.921	0.924	-0.002	-0.002	-0.010	-0.010	0.000	0.002
c = 0.1, r = 7.5											
q_{10}	3.192	3.982	2.941	3.129	2.888	0.251	0.241	0.790	0.800	0.010	0.037
q_{50}	13.459	18.460	12.808	12.809	12.181	0.651	0.628	5.000	5.023	0.023	0.114
q_{90}	56.775	68.385	55.771	52.450	51.370	1.004	1.080	11.610	11.534	-0.076	0.646
$F_{Y\mid DT}(q_{10\mid 11})$	0.140	0.100	0.156	0.145	0.161	-0.016	-0.016	-0.040	-0.039	0.000	0.002
$F_{Y\mid DT}(q_{50\mid 11})$	0.610	0.500	0.623	0.629	0.643	-0.014	-0.014	-0.110	-0.110	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.927	0.900	0.928	0.936	0.938	-0.002	-0.002	-0.027	-0.027	0.000	0.002
c = 0.5, r = 2.5											
q_{10}	4.762	5.025	3.085	4.174	2.888	1.677	1.286	0.263	0.654	0.391	0.393
q_{50}	20.074	21.741	15.641	15.644	12.181	4.433	3.463	1.667	2.637	0.970	0.981
q_{90}	84.646	88.516	79.253	58.641	51.370	5.393	7.270	3.870	1.993	-1.877	2.066
$F_{Y\mid DT}(q_{10\mid 11})$	0.109	0.100	0.190	0.135	0.216	-0.081	-0.081	-0.009	-0.009	0.000	0.002
$F_{Y\mid DT}(q_{50\mid 11})$	0.528	0.500	0.598	0.625	0.696	-0.070	-0.071	-0.028	-0.028	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.907	0.900	0.916	0.954	0.962	-0.009	-0.009	-0.007	-0.007	0.000	0.002
c = 0.5, r = 7.5											
q_{10}	4.762	5.552	3.085	4.174	2.888	1.677	1.286	0.790	1.181	0.391	0.393
q_{50}	20.074	25.074	15.641	15.644	12.181	4.433	3.463	5.000	5.970	0.970	0.981
q_{90}	84.646	96.256	79.253	58.641	51.370	5.393	7.270	11.610	9.733	-1.877	2.066
$F_{Y\mid DT}(q_{10\mid 11})$	0.127	0.100	0.212	0.158	0.243	-0.086	-0.086	-0.027	-0.026	0.000	0.002
$F_{Y\mid DT}(q_{50\mid 11})$	0.578	0.500	0.640	0.676	0.738	-0.062	-0.062	-0.078	-0.078	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.919	0.900	0.926	0.961	0.968	-0.007	-0.007	-0.019	-0.019	0.000	0.001
c = 0.75, r = 2.5											
q_{10}	6.114	6.377	3.133	4.874	2.888	2.980	1.986	0.263	1.257	0.994	0.995
q_{50}	25.769	27.435	17.715	17.728	12.181	8.053	5.547	1.667	4.173	2.507	2.513
q_{90}	108.525	112.395	100.071	64.488	51.370	8.454	13.118	3.870	-0.794	-4.664	4.784
$F_{Y\mid DT}(q_{10\mid 11})$	0.107	0.100	0.235	0.155	0.283	-0.128	-0.128	-0.007	-0.007	0.000	0.002
$F_{Y\mid DT}(q_{50\mid 11})$	0.522	0.500	0.618	0.668	0.764	-0.096	-0.096	-0.022	-0.022	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.905	0.900	0.916	0.967	0.977	-0.010	-0.010	-0.005	-0.005	0.000	0.001
c = 0.75, r = 7.5											
q_{10}	6.114	6.904	3.133	4.874	2.888	2.980	1.986	0.790	1.784	0.994	0.995
q_{50}	25.769	30.769	17.715	17.728	12.181	8.053	5.547	5.000	7.507	2.507	2.513
q_{90}	108.525	120.135	100.071	64.488	51.370	8.454	13.118	11.610	6.946	-4.664	4.784
$F_{Y\mid DT}(q_{10\mid 11})$	0.121	0.100	0.254	0.175	0.308	-0.133	-0.133	-0.021	-0.021	0.000	0.003
$F_{Y\mid DT}(q_{50\mid 11})$	0.562	0.500	0.648	0.708	0.794	-0.086	-0.086	-0.062	-0.062	0.000	0.003
$F_{Y\mid DT}(q_{90\mid 11})$	0.915	0.900	0.924	0.971	0.980	-0.009	-0.009	-0.015	-0.015	0.000	0.001

Notes: The data are generated according to the DGP as defined in (3.26) and (3.27), with $N = 400{,}000$, sampled 500 times. Hence, indicated values in columns 2-8 refer to averages over all 500 MC repetitions. The table contains statistics for the 10th, 50th, and 90th percentiles, as well as for the population shares evaluated at the relevant quantile for the QTT, i.e., $q_{\tau\mid 11}$, of each of the four sub-populations, including the counterfactual values for the sub-population $(D,T) = (1,1)$ in column 2. The temporal differences in potential outcomes for the treated observations and the control group observations are displayed for each statistic. Furthermore, the true (oracle) treatment effects on the treated for statistic φ, $\psi_0(\varphi)$, alongside a non-parametric double-difference DiD estimator, $\widehat{\psi}_{DiD}(\varphi)$, are displayed, with φ referring to either the quantile or population share, respectively (see equations (3.23) and (3.24)). Additionally, the last two columns refer to the estimated $Bias = (1/M)\sum_m[\widehat{\psi}_{m,DiD}(\varphi) - \psi_{m,0}(\varphi)]$, and the $RMSE = [(1/M)\sum_m(\widehat{\psi}_{m,DiD}(\varphi) - \psi_{m,0}(\varphi))^2]^{(1/2)}$, with $m = 1,\ldots,500$. $\psi_{m,0}(\varphi)$ is the true oracle treatment effect in the mth MC repetition. Analogously, $\widehat{\psi}_{m,DiD}$ denotes the DiD estimation obtained in the mth MC repetition. Note that the oracle effects for quantile effects can be exactly computed as $\psi_0(q_\tau) = QTT_\tau = r \times [\tau + (1/2)\cdot\tau^2 + (1/3)\cdot\tau^3]$, $\tau \in [0,1]$, i.e., $\psi_{m,0}(q_\tau) = \psi_0(q_\tau) = QTT(\tau)$. This is not the case for the population share effect, $\psi_{m,0}(F_{Y\mid DT}(q_{\tau\mid 11}))$, since the DGP is characterized by the sum of two lognormal distributions. Therefore, the treatment effect does not have a closed-form analytic solution that allows for the exact determination of values. Following Firpo and Pinto (2016, p.473), values of ψ_0 displayed in column 9 that refer to population shares thus refer to the median value across the 500 MC simulations.

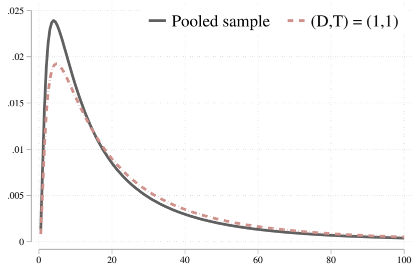
Figure 3.5: Comparison of pooled and marginal $(D,T) = (1,1)$ distributions for varying common trend and treatment intensities



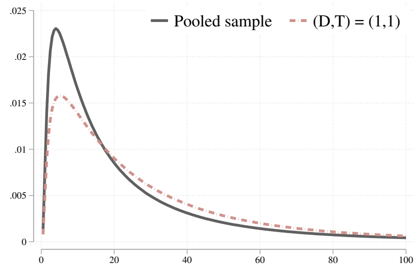
$c = 0.1$, $r = 2.5$



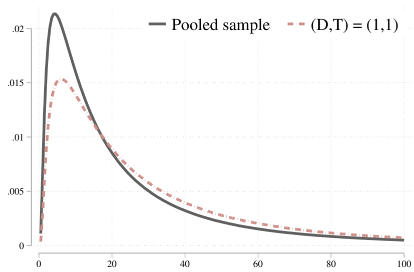
$c = 0.1$, $r = 7.5$



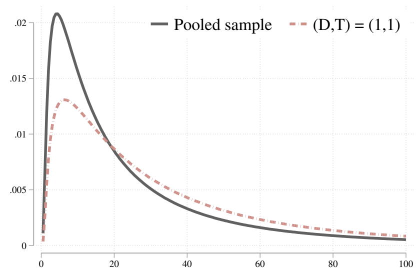
$c = 0.25$, $r = 2.5$



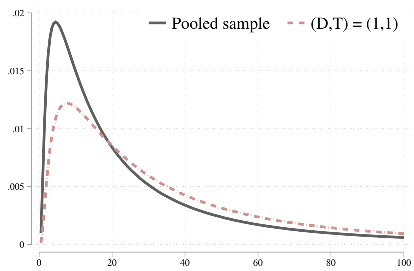
$c = 0.25$, $r = 7.5$



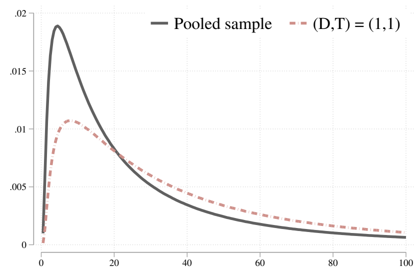
$c = 0.5$, $r = 2.5$



$c = 0.5$, $r = 7.5$



$c = 0.75$, $r = 2.5$



$c = 0.75$, $r = 7.5$

Notes: Figures show the comparison between the marginal distribution of the $(D,T) = (1,1)$ sub-population and the distribution of the pooled sample including all four sub-populations with varying common trend and treatment intensities. All figures are based on simulations with $N = 100,000$ per sub-population that are sampled 500 times. The displayed distributional masses represent the average over all 500 MC samples. Figures are based on the DGP as specified in (3.26) and (3.27). Pdfs are computed using a bin-wise consideration, i.e., the distributional mass for a given threshold, z_j, is given by $\mathbb{P}[Y \leq z_j] - \mathbb{P}[Y \leq z_{j-1}]$, where the increments between thresholds are 0.5.

Since the DGP implies a common trend in population shares but not in quantiles, the bias and the RMSE increase with c but not with r for DiD estimators that concern quantiles. Crucially, no bias and a RMSE close to zero are detectable for population share DiD estimates along either c or r.

What does these results imply for the RIF-based methods? First and foremost, observable differences between the DDiD- and the QDiD-RIF approach are expected to be shaped by the different identifying assumptions upon which they are based. The more general findings obtained for the non-parametric double-difference DiD estimator in table 3.1 will thus help rationalize the differences between these two RIF-DiD methods.

However, these results are not informative regarding the bias of a pooled RIF approach, as biased estimates arise from the misleading underlying distribution used in the QTT approximation. Hence, to gain a better understanding of the sources of bias in that case, figure 3.5 depicts the outcome distributions of both the pooled sample and the marginal $(D,T) = (1,1)$ sub-population. The two distributions appear similar for small values of c and r, but the differences become substantial as c and/or r increase. Since the increasing disparity between the pooled and marginal distributions leads to biased results when using the pooled RIF approach, both the bias and RMSE of the pooled RIF-DiD estimator are expected to increase with c and r.

Simulation results: RIF-based estimation of QTTs

In the following, the performance of the pooled RIF, the DDiD-RIF, and the QDiD-RIF approach are examined. Since both $F_{Y_i\,|\,11}$ and $F_{Y_i^0\,|\,11}$ are known, it is possible to determine the true oracle QTTs and compare them with the estimated QTTs obtained from all three approaches. Figure 3.6 and figure 3.7 illustrate the bias and the RMSE, respectively, across various c/r combinations. Moreover, figure 3.8 displays the comparison between the estimated QTTs and the true target values.[22]

The first column of figure 3.6 confirms the ex ante conjectures regarding the pooled RIF approach's responsiveness of the bias that becomes more pronounced as both common trend and treatment intensity increases. Noticeably, the bias follows a U-shaped pattern, implying a negative bias that is largest around the 45th/50th percentile and increases be-

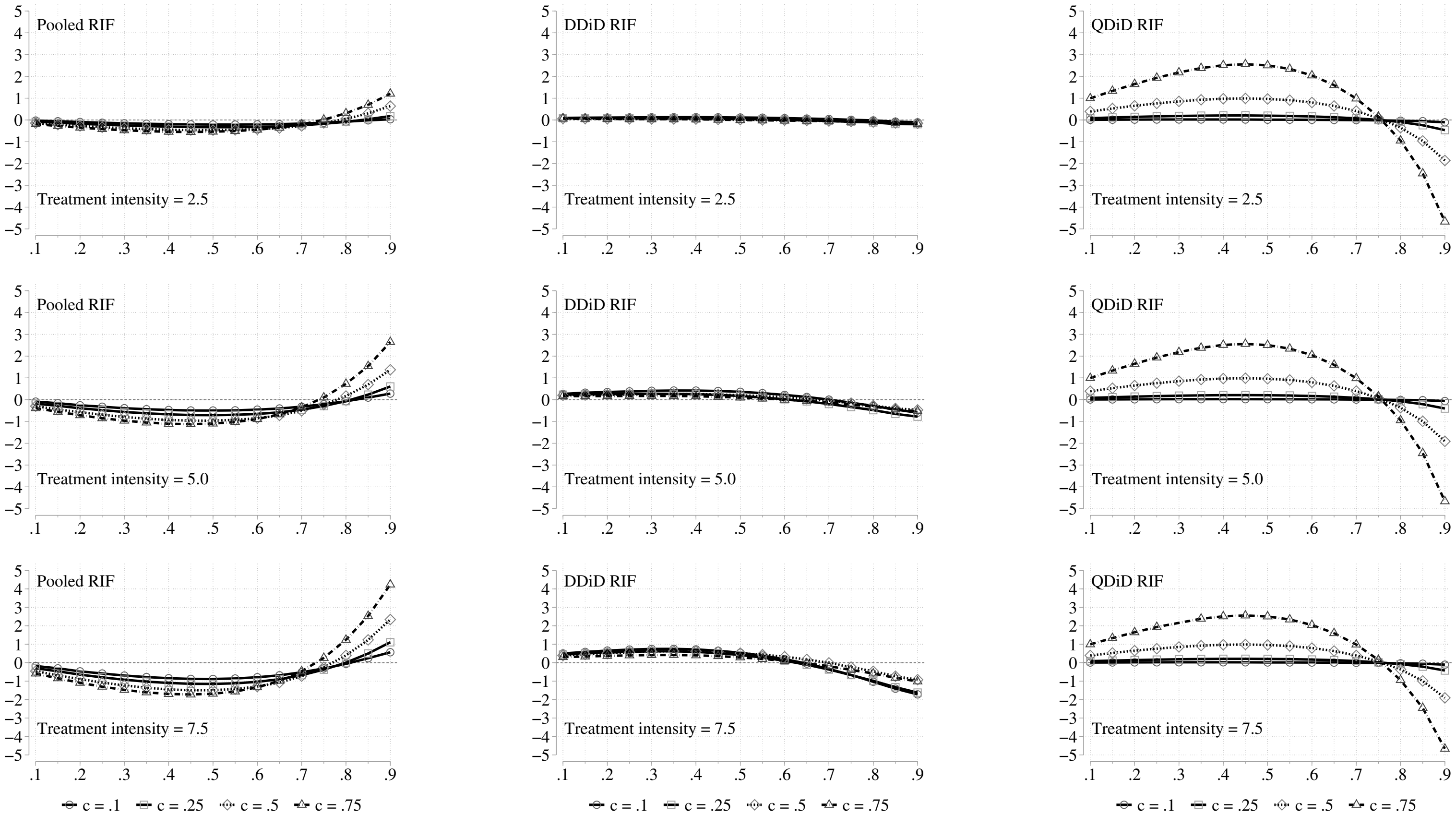
[22] For the full MC simulation results see appendix B.3.1.

yond this point. This pattern arises due to the misleading approximation inherent in the pooled RIF-DiD approach, which is manifested in the division by the slope of the cdf in equation (3.16). Specifically, the DGP design imposes a rightward shift in the distribution due to both the common trend and the treatment effect. Considering figure 3.5, this implies that – in comparison to the marginal $(D,T)=(1,1)$ outcome distribution – the slopes of the pooled sample cdf are excessively steep for lower values and overly flat for upper values of the outcome distribution. Consequently, the pooled RIF-DiD approximation in equation (3.16) employs a slope for the first-order approximation that is "too large" for values up to about 20 and "too small" for values beyond 20, with values around 20 roughly coinciding with the 45th/50th percentiles of the $(D,T)=(1,1)$ marginal outcome. While this induces a downward bias for quantiles below the 45th/50th percentile, the differences in the slopes reverse sign above this point (cf., figure 3.5), i.e., the slope of the pooled cdf becomes "too flat". This, in turn, explains the upward bias at higher quantiles.

Moreover, because the pooled distribution is used instead of the marginal $(D,T)=(1,1)$ distribution, the quantiles around which the approximations are made are misleading. Specifically, the difference between the pooled and the marginal $(D,T)=(1,1)$ distribution implies that $q^m_\tau < q_{\tau|11}$, or equivalently, $q^m_\tau = q_{\check{\tau}|11}$ with $\check{\tau} < \tau$. As the treatment effect is defined such that the QTT_τ increases with τ (cf., equation (3.27)), this indicates that the measured treatment effect in the numerator of (3.16) must lag behind the true treatment effect (which, at q^m_τ, is determined by $\check{\tau}$). Overall, the findings in figure 3.6 show that this lag contributes to the downward bias at lower quantiles, which is overshadowed by the upward bias caused by the too-flat slope at higher quantiles.

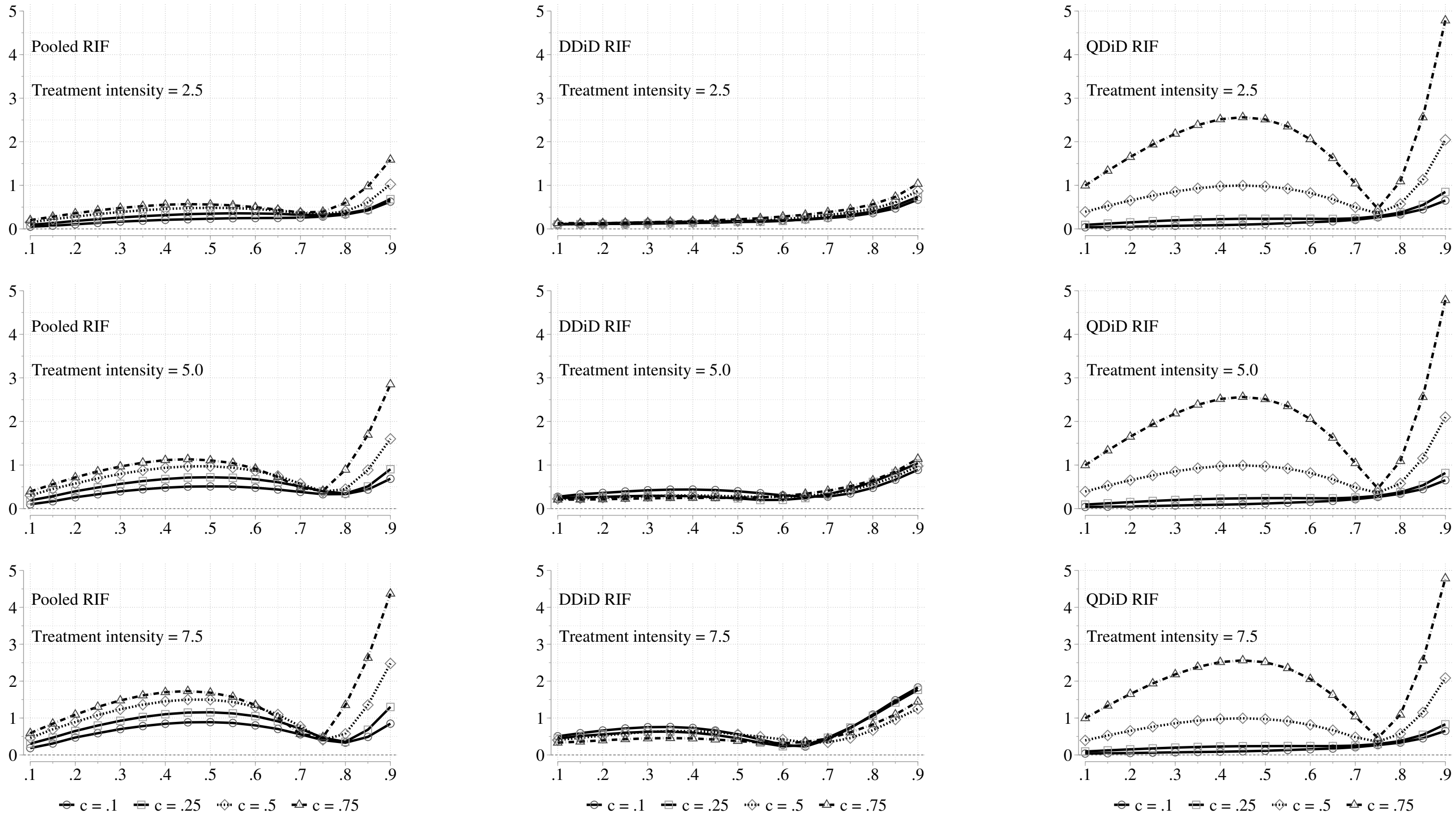
The bias of the DDiD-RIF estimates is presented in the second column of figure 3.6. As previously hypothesized, the common trend intensity impacts the bias virtually imperceptibly, resulting in a nearly negligible implied bias for small and medium treatment intensities. Not surprisingly, the approximation of the actual effect is still associated with a non-negligible bias for larger treatment intensities, mirroring the approximative shortcoming inherent to a RIF approach. Formally, the only source of bias that can explain the reverse U-shape pattern for large treatment intensities is the cdf's slope in (3.19), $f_{Y|11}(q_{\tau|11})$, reflecting the approximation procedure. On the contrary, the population share effect, represented by γ^{DR}_τ, is unbiased, as shown by the results in table 3.1 for the DGP considered herein.

Figure 3.6: Bias comparison along varying common trend and treatment intensities for the pooled RIF, DDiD-RIF, and the QDiD-RIF approach



Notes: The figures show the bias of all estimates obtained from each of the three RIF approaches for different quantiles with different common trend/treatment intensities. The horizontal axis refers to the τth quantile considered. All figures are based on simulations of the DGP as specified in (3.26) and (3.27), with $N = 400,000$ sampled $M = 500$ times. The depicted bias is computed as $(1/M)\sum_m \widehat{QTT^v}_{\tau,m} - QTT_\tau$, with $v \in \{$pooled, DDiD, QDiD$\}$. The oracle QTT is given as $QTT_\tau = r \times [\tau + (1/2)\cdot\tau^2 + (1/3)\cdot\tau^3]$, $\tau \in (0,1)$. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), or (3.22), respectively.

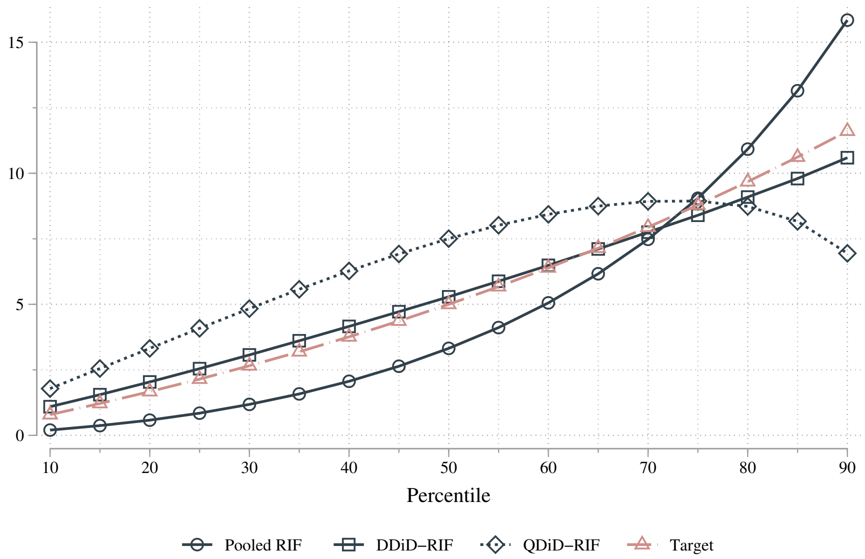
Figure 3.7: RMSE comparison along varying common trend and treatment intensities for the pooled RIF, DDiD-RIF, and the QDiD-RIF approach



Notes: The figures show the RMSE of all estimates obtained from each of the three RIF approaches for different quantiles with different common trend/treatment intensities. The horizontal axis refers to the τth quantile considered. All figures are based on simulations of the DGP as specified in (3.26) and (3.27), with $N = 400,000$ sampled $M = 500$ times. The depicted RMSE is computed as $[(1/M)\sum_m(\widehat{QTT^v}_{\tau,m} - QTT_\tau)^2]^{(1/2)}$, with $v \in \{\text{pooled, DDiD, QDiD}\}$ The oracle QTT is given as $QTT_\tau = r \times [\tau + (1/2) \cdot \tau^2 + (1/3) \cdot \tau^3]$, $\tau \in (0,1)$. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

The implied bias of the QTT estimate, obtained by means of the DDiD-RIF approach, suggests an overly large approximation at lower to middle quantiles, decreasing towards the 60th percentile, after which the bias turns negative. This pattern can be explained by the RIF being a *linear* approximation. Considering figure 3.8, the pattern of the DDiD-RIF estimates is nearly perfectly linear across quantiles, underscoring the capability of any RIF approach to accurately approximate location shifts.

Figure 3.8: Comparison of QTT estimates of the pooled RIF, DDiD-RIF, and the QDiD-RIF approach



Notes: The figure shows the coefficients alongside the true target values for the respective QTT. By way of example, the target and the respective coefficients for $c = 0.75$ and $r = 7.5$ are displayed to make the involved patterns more salient. All figures are based on simulations with $N = 400,000$ that are sampled 500 times. The target values are the true QTTs for $r = 7.5$, i.e., $Target = 7.5 \times [\tau + (1/2) \cdot \tau^2 + (1/3) \cdot \tau^3]$. The coefficients for the displayed approaches refer to the average values over all the 500 MC samples. Figures are based on the DGP as specified in (3.26) and (3.27).

On the other hand, the true target effect follows a convex pattern due to the non-linear setting of the treatment effect in (3.27) that implies a different shape of the cdf as well. Taken together, this ‘secant-behavior’ that is apparent for the DDiD-RIF QTT approximations in figure 3.8 explains the over-/under-estimation of QTTs below/above the 60th percentile. As to the increasing bias at higher quantiles, recall that the treatment effect in (3.27) is modeled as rank-dependent, i.e., the non-linear elements (R_i^2 and R_i^3) are larger at those upper quantiles. This, in turn, explains the comparatively large bias for upper quantiles.

Lastly, the third column of figure 3.6 shows the bias of estimates obtained using the QDiD-RIF approach. As shown in table 3.1, the bias in quantile DiD estimates, which rely on the QDiD assumptions, is entirely driven by the intensity of the common trend if this assumption is not met. Consequently, as shown in figure 3.6, the bias remains unchanged across different treatment intensities but increases only with the common trend intensity. Specifically, the simulation results indicate that the QDiD-RIF estimates imply counterfactual quantiles that fall short of the true value up to the 75th percentile, leading to an overestimation of the effect. Beyond this point, the relationship reverses, indicating a negative bias for percentiles above the 75th percentile. This pattern naturally becomes more pronounced with higher common trend intensities. Moreover, the fact that the bias is quite small for low common trend intensities is due to the fact that for these cases the DDiD and QDiD assumptions become similar, as formally outlined in appendix B.1.1.

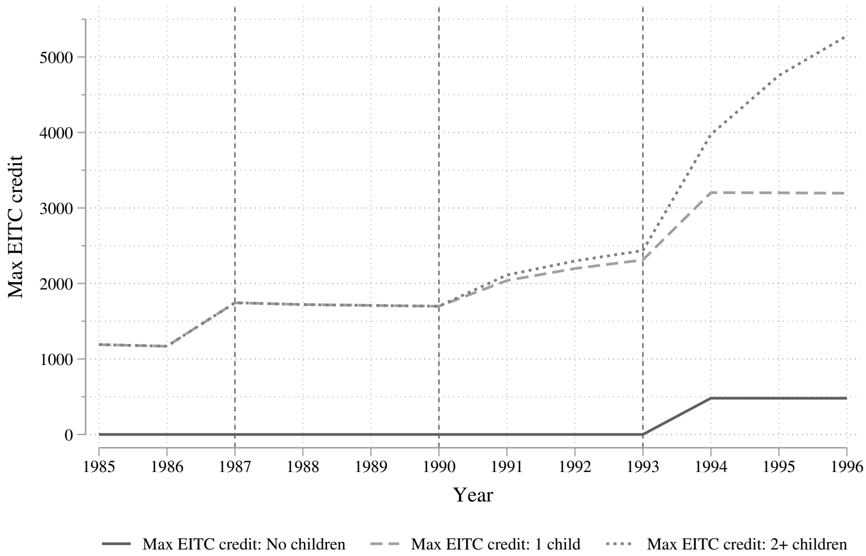
Finally, the comparison of the RMSEs in figure 3.7 reflects the findings from figure 3.6, indicating that for the given DGP, the DDiD-RIF approach is best suited for estimating the QTT when there is a non-negligible common trend. In contrast, the QDiD-RIF approach provides reliable results only for small common trend intensities. Moreover, it is crucial to note that for the DGP considered, the pooled RIF approach is consistently outperformed by the DDiD-RIF approach in terms of RMSE, even when the distributions of sub-populations differ only moderately. Furthermore, the slightly higher RMSE of the DDiD-RIF approach at larger treatment intensities is due to its reliance solely on the $(D,T) = (1,1)$ sub-population to estimate the cdf's slope, leading to greater variance in the effects compared to the pooled RIF approach. The pooled RIF approach, on the other hand, utilizes the entire pooled distribution, which in this case, is three times larger.

3.6 Empirical illustration

To illustrate the performance of all three RIF-based approaches, we examine the distributional impact of the EITC on income-to-poverty thresholds for low-income single mothers, following the approach of Hoynes and Patel (2018).

The literature on the effect of the EITC or expansions thereof is substantial, including contributions that examine the effect on, e.g., labor force participation and earnings.[23] In short, the EITC is an important transfer program in the United States that targets low-income families with children. Taxpayers are eligible for the EITC if they earned income in a given tax year and had no more than a specified amount of 'adjusted gross income and earned income', with steeply increasing credits per dollar earned at low income levels (cf., Hoynes and Patel, 2018, pp. 862/863). Moreover, the EITC is designed such that low-income families with children are the main beneficiaries of the policy, constituting a comparatively successful redistributional tool in the US (Hoynes and Rothstein, 2016).

Figure 3.9: Evolution of maximum EITC credit



Notes: Schedule of the maximum EITC that could be collected by families with 0, 1, or 2+ children. Vertical dashed lines indicate the occurrence of EITC reforms as described in detail in Hoynes and Patel (2018).
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this graph are part of the replication data provided in Hoynes and Patel (2022).

To examine the distributional effect of one of the largest extension of the EITC in 1993 ('OBRA93', cf., figure 3.9), Hoynes and Patel (2018) consider the distribution of multiples of official poverty thresholds as their outcome variable. This expansion, and the fact that the benefits disproportionately accrued to families with children, constitutes the quasi-experimental setting exploited within a DiD framework by Hoynes and Patel

[23] See, for example, Nichols and Rothstein (2016) for an in-depth examination of the EITC.

(2018). In this study, single women aged 24 to 48 with some college education were used as the relevant sample.

The approach by Hoynes and Patel (2018) involves a distributional DiD analysis examining the policy's effect on the propensity to fall above multiples of their dependent variable. Alternatively, instead of considering the effects on these propensities, an approach that allows to examine quantile effects would have also been a viable option. Since the actual value of the outcome variable is of interest here, this procedure implies the necessity to backwardly infer the value of the outcome variable from an estimated quantile treatment effect to properly interpret the findings at a given position of the unconditional outcome distribution. Note that this procedure of 'backwardly inferring' the effect around given values of the outcome variable's support has been used in previous contributions as well, highlighting the importance of a sound understanding around which value the effect is actually estimated.[24]

In the subsequent analysis, each of the three aforementioned RIF approaches is employed, allowing for a comparison of the estimated effects. Moreover, the results are compared with the findings in Hoynes and Patel (2018) to assess the plausibility of the RIF-based approaches. Unlike the unconditional case discussed earlier, a comprehensive set of control variables is considered, as recommended in Hoynes and Patel (2018), when estimating the QTTs. To that end, the edited and cleaned data provided by Hoynes and Patel (2022) are used.[25]

In the following, the 2×2 DiD specification considered by Hoynes and Patel (2018) to evaluate the OBRA93 expansion on the distribution of income-to-poverty thresholds is reexamined, with the treatment being defined as 'having one or more children'. As in

[24]The following two examples 'backwardly infer' results from RIF-DiD results. On the one hand, the RIF-DiD approach used in the study by Bossler and Schank (2023, see detailed review above) entails a 'backward inference' of effects. When interpreting their effect sizes against the background of a special form of employment in the German system, Bossler and Schank (2023, p. 833) note that "[T]his is because the 10th percentile is located right at the minijob threshold (...)", with the minijob threshold amounting to 450 euros. On the other hand, Bossler et al. (2024, p. 18) also backwardly infer the effect of the German minimum wage introduction on a specific hours value from their RIF-regression specification: "The negative hours effect at the 16th percentile is observed at a point of the hours distribution where employees are working just above 50 hours per month."

[25]The corresponding replication do files for the data preparation/cleaning/analyses, as well as the raw and edited data sets were made available by Hoynes and Patel (2022) for replication and can be accessed online: `https://doi.org/10.7910/DVN/TCUGHY`.

Hoynes and Patel (2018), the pre-treatment period comprises the years 1992 and 1993, whereas the post-treatment period involves the years from 1994 to 1999. The empirical specification specified by Hoynes and Patel (2018) reads

$$\mathbb{1}[Y_{it} > k] = \alpha + \beta(post_t \times treat_c) + \eta_{st} + \gamma_c + X_{it}\Phi + \varepsilon_{it}, \tag{3.28}$$

with i referring to individual taxpayers, t being the tax year, η_{st} capturing a set of state by year fixed effects, γ_c are dummy variables indicating the actual number of children (0,1,2,3+), and k referring to various income-to-poverty thresholds. The vector X_{it} comprises additional socioeconomic characteristics as described in detail in Hoynes and Patel (2018). The treatment effect, β, hence corresponds to the effect of the OBRA93 expansion on the propensity to fall above the kth income-to-poverty thresholds for families with at least 1 child ($treat_c = 1$) in the years after the expansion ($post_t = 1$).

Instead of constructing several indicator variables for values falling above or below certain thresholds, the outlined RIF-based approaches are employed. Firstly, this involves forming the RIF according to (3.15) and running the regression as in (3.28) for the *pooled RIF* approach. As a second approach, QTTs are estimated employing the *DDiD-RIF* approach as in (3.20). Lastly, QTTs are estimated by means of running the regression as in (3.28) with the compound RIF according to (3.21) as the dependent variable for the *QDiD-RIF* approach.[26]

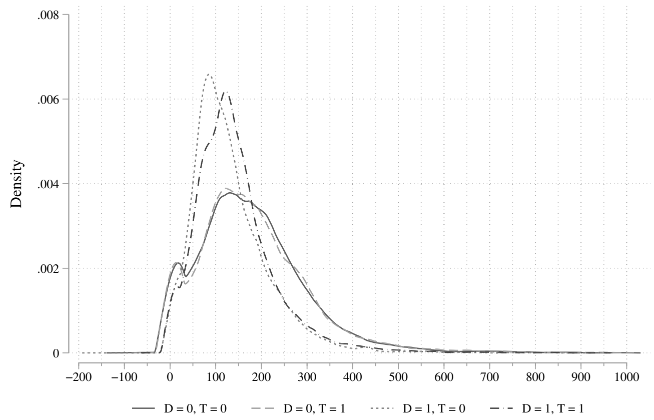
Before comparing the estimates of the three approaches, consider figure 3.10a displaying the distribution of all four involved sub-populations. The two groups are markedly different, with the control group being characterized by more distributional mass at upper income-to-poverty thresholds, while the treatment group distribution (single mothers with at least one children) is concentrated around the threshold of 100. Moreover, it appears that the OBRA93 expansion had an effect on the treated group but not on the control group, with virtually no shifts being observable from the pre- to the post-treatment period for the control group and non-negligible shifts for the treatment group. Furthermore, figure 3.10b illustrates the disparity between the marginal distribution for the treated group in

[26] Note that the DiD analysis by Hoynes and Patel (2018) imply an assumed common trend assumption in population shares (the DDiD assumptions from above), and hence the appropriate RIF approach would in fact be the DDiD-RIF approach. However, as another goal of the application is to single out the differences between the two proposed approaches vis-à-vis the pooled RIF approach, the QDiD-RIF approach is considered here as well.

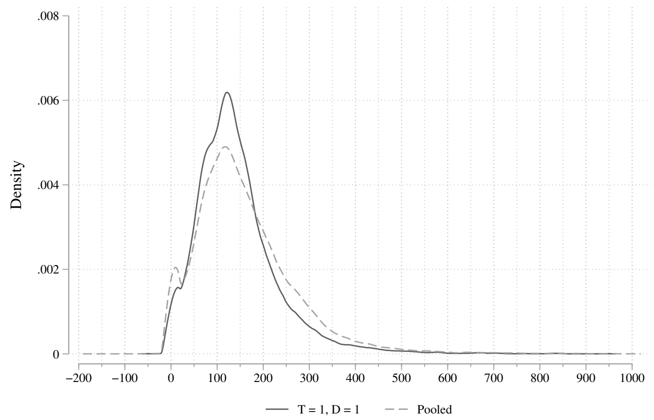
the post-treatment period and the pooled distribution, comprising all four sub-populations.

Figure 3.10: Comparison of outcome distributions

(a) Comparison of involved sub-populations



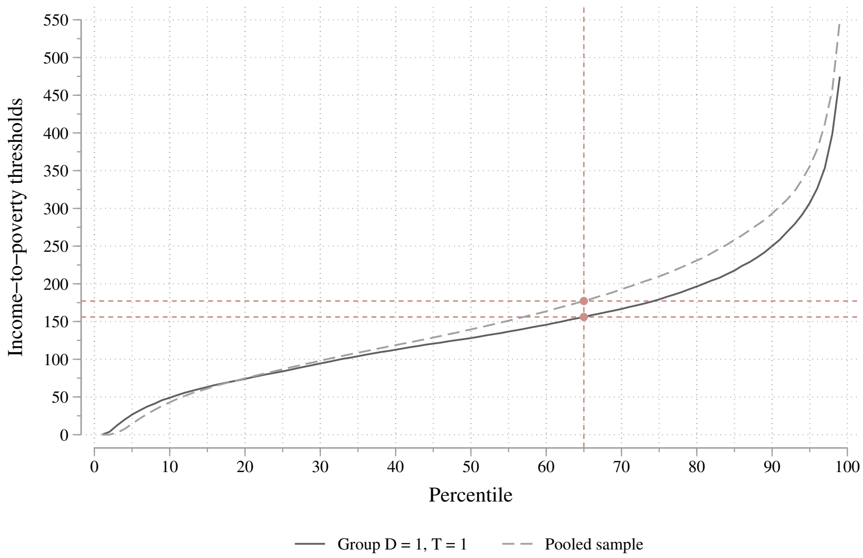
(b) Comparison of pooled distribution vs. marginal distribution in $(D,T) = (1,1)$



Notes: Comparison of involved distributions in various RIF-based approaches: Subfigure 3.10a shows the income-to-poverty outcome distribution for all four sub-populations. Subfigure 3.10b shows the comparison of the pooled and marginal distributions for the treated group in the post-treatment period.
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this graph are part of the replication data provided in Hoynes and Patel (2022).

The most noticeable distinction is that the pooled distribution has more mass at higher thresholds and less at lower-middle thresholds compared to the marginal $(D,T) = (1,1)$ outcome distribution. To further emphasize the differences between the pooled and marginal distributions, figure 3.11 depicts the quantile functions derived from both distributions. As an illustrative example, consider the 65th percentile of income-to-poverty thresholds derived from the *pooled* distribution. Compare this with the corresponding values derived from the *marginal* distribution of the post-treatment/treated sub-population. This 65th percentile corresponds to an income-to-poverty threshold value of 177 when considering the *pooled* distribution, whereas the corresponding value for this percentile in the *marginal* treated/post-treatment sub-population is 156. This difference is important for interpreting the RIF-based estimates due to the need to subsequently infer the values around which approximations were made.

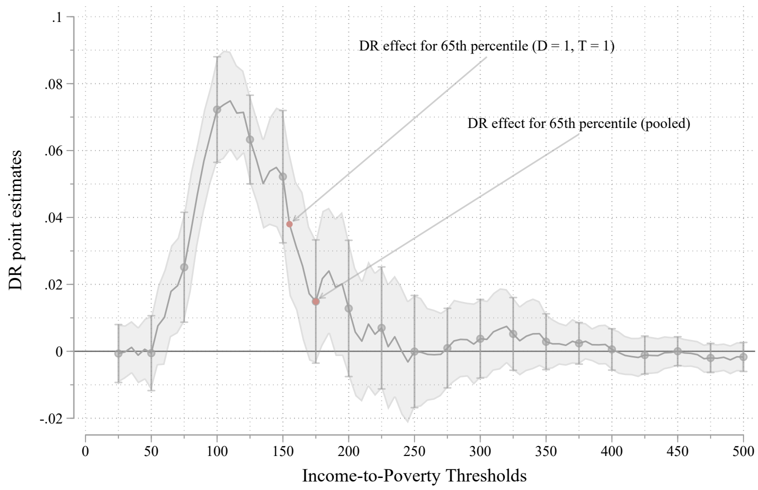
Figure 3.11: Comparison of quantile functions



Notes: The figure depicts quantile functions derived from the *pooled* and the *marginal* $(D,T) = (1,1)$ distribution. To highlight the differences between the two, the income-to-poverty thresholds corresponding to the *65th percentile* of each distribution are marked by the red vertical line. For the pooled distribution, the *65th percentile* corresponds to an income-to-poverty threshold of *177.179*. In the marginal $(D,T) = (1,1)$ distribution, the equivalent threshold is *156.022*, as shown by the intersection with the lower red horizontal line.
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this graph are part of the replication data provided in Hoynes and Patel (2022).

Given these disparities, it becomes apparent that acknowledging the underlying values around which the effects are to be approximated is crucial. This is further demonstrated by the illustration of the extended results in Hoynes and Patel (2018, Figure 6, p. 879) in figure 3.12, which both replicates and augments their findings by means of additional estimates of the effects for the thresholds in-between their 25%-point increments. To collect some results that can be used to compare the results from the RIF approaches, first consider the effect for the threshold closest to the value 156 – the 65th percentile of the *marginal* distribution. The distribution regression results for this value indicate a statistically significant treatment effect on the treated, corresponding to an increase in the propensity to exceed this threshold by approximately 4 percentage points.

Figure 3.12: Replication of Hoynes and Patel (2018, Figure 6)



Notes: Replication of the distribution regression (DR) results from Hoynes and Patel (2018, Figure 6). Estimates refer to point estimates of β in (3.28). Scatters and whiskers highlight the results presented in their paper. The solid line represents findings for additional thresholds with increments of 5 (i.e., $25, 30, 35, \ldots, 495, 500$). The grey shaded area corresponds to the point-wise 95% confidence interval, clustered at the state level. To compare the DR findings with those from the RIF approaches, the results for the thresholds closest to the 65th quantile of the pooled (177 – corresponding highlighted threshold: 175) and the marginal $(D = 1, T = 1)$ distribution (156 – corresponding highlighted threshold: 155) are highlighted in red.
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this graph are part of the replication data provided in Hoynes and Patel (2022).

On the other hand, results around the threshold closest to 177 (the 65th percentile of the *pooled* distribution) suggest an insignificant effect that is also much smaller in magnitude. Based on these distribution regression findings, a valid RIF approach (i.e., DDiD- or QDiD-RIF) should indicate non-negligible shifts around the 65th percentile. Contrarily, the pooled RIF approach is expected to underestimate the QTT for the 65th percentile. This is because the effect would be approximated around a misleadingly large – i.e., too large – income-to-poverty threshold value around which only a small and insignificant shift in distributional masses, as implied by the treatment, was estimated.

Estimation results: RIF-DiD approaches

Figure 3.13 shows QTT estimates obtained using the pooled, DDiD-, and QDiD-RIF approaches. To examine how the use of the pooled distribution contributes to the over- or underestimation of QTTs when employing a pooled RIF-DiD approach, the presentation in figure 3.13 explicitly contrasts the pooled RIF-DiD findings with those of the other two approaches. Furthermore, to ascertain which values of the outcome distribution the respective QTT estimates correspond to, the percentiles and the underlying thresholds from both the pooled and marginal outcome distributions are indicated.

Following the example from above, the QTT estimate for the 65th percentile is examined in greater detail to highlight the differences between the approaches. Using the pooled RIF-DiD approach, the corresponding QTT estimate suggests an increase of approximately 5 in the income-to-poverty threshold, with the effect approximated around a value of 177. However, this effect is not statistically significant at the 5% level. Recall that RIF-based approaches necessitate inferring the values of the outcome distribution from the QTTs retrospectively to interpret the quantile-based effects, as previously described. This indicates that the pooled RIF-DiD estimate is misleading because it does not correspond to the actual quantile of interest – the relevant value for the 65th percentile of the income-to-poverty threshold should be 156, not 177.

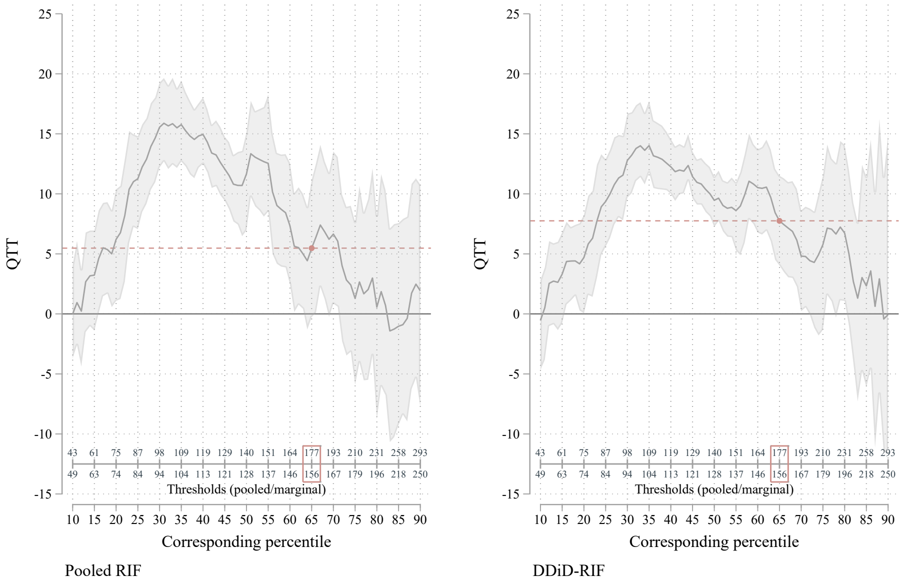
Using either the DDiD- or QDiD-RIF approach, which ensures an approximation around the correct threshold, reveals substantially larger effects that are statistically significant. Specifically, the QTT estimates obtained by means of the DDiD- (QDiD-)RIF approach indicate a statistically significant increase of 7.8 (8.6) in the income-to-poverty threshold, with effects being approximated around the 65th percentile of the marginal $(D,T) = (1,1)$ outcome distribution. Put differently, the QTT estimate for the 65th percentile using the misleading pooled RIF-DiD approach underestimates the QTT by a magnitude of 2.3 (3.1) or 30% (36%) when compared with the DDiD- (QDiD-) RIF approach.[27]

Overall, there are limited differences between the pooled RIF-DiD and the DDiD- and QDiD-RIF approaches up to the 50th percentile, but effects are smaller and tend to be statistically insignificant beyond that point, standing in contrast to the DDiD- and QDiD-RIF approaches. Reconsidering the quantile functions depicted in figure 3.11, finding similar values for the lower percentiles is not surprising, given that the pooled and the marginal $(D,T) = (1,1)$ distribution appear quite similar up to roughly the 40th percentile and only start to diverge significantly from the median onward.

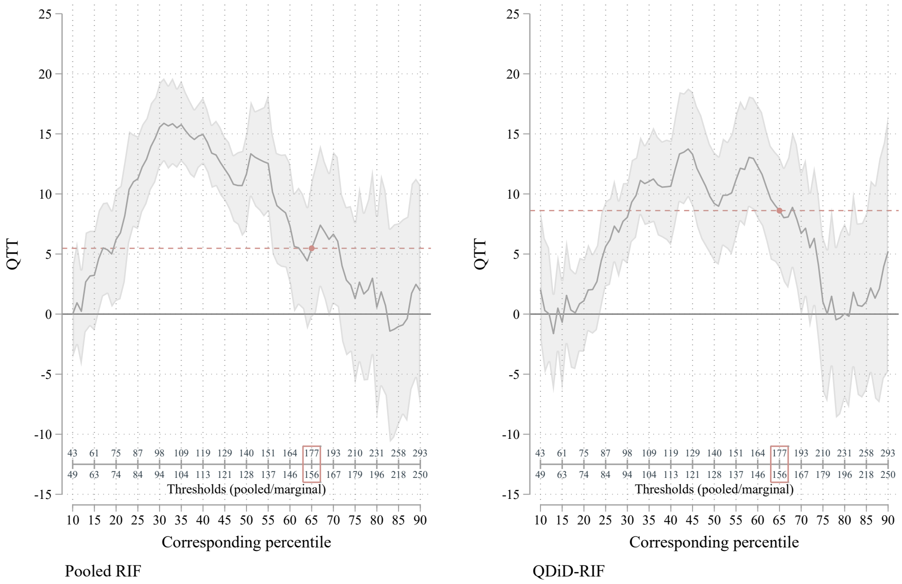
[27] For full results, see table B.9.

Figure 3.13: Comparison of QTT estimates using different approaches

(a) Pooled RIF DiD vs. DDiD-RIF



(b) Pooled RIF DiD vs. QDiD-RIF



Notes: Comparison of QTT estimates of the OBRA93 expansion's effect on percentiles of the income-to-poverty outcome distribution. Subfigures 3.13a and 3.13b show the comparison of estimates between the pooled RIF-DiD approach and the DDiD-/QDiD-RIF approaches. The red horizontal lines highlight the QTT estimate for the 65th percentile of the respective RIF approach. The superimposed x-axis depicts the corresponding threshold values for the τth percentiles with $\tau \in 10, 15, \ldots, 90$ of both the pooled (upper values) and the marginal distribution (lower values). The threshold values corresponding to the 65th percentile of both the pooled and marginal distributions are highlighted by a red box on the superimposed x-axis.
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this graph are part of the replication data provided in Hoynes and Patel (2022).

3.7 Conclusion

Several recent studies investigating the distributional treatment effects of policies, such as minimum wages, have combined unconditional quantile regressions in the form of RIF regressions with a linear DiD approach to estimate the quantile treatment effect on the treated. This paper examines the appropriateness of RIF-DiD approaches for estimating QTTs. Considering the simple 2×2 DiD case, it is shown that issues with a 'pooled RIF-DiD' approach – which uses all four sub-populations to estimate QTTs – stem not only from the inherent approximative shortcomings of the RIF but also from structural shortcomings. This makes such an approach generally ill-suited for estimating QTTs. The main error emerges from approximating effects around wrong quantiles which leads to biased estimates. This erroneous structure becomes more aggravated as the differences between the pooled and the relevant marginal outcome distributions increase, i.e., as common trend and/or treatment intensities grow larger. The present paper has demonstrated this issue both formally and by means of extensive Monte Carlo simulation results.

Two solutions to overcome the structural issue are proposed. First, an easily implementable DDiD-RIF approach is examined in greater detail. The difference between this and the pooled RIF-DiD approach is that the effect is approximated around the actual quantile of interest – specifically, the quantile of the treated group in the post-treatment period. As a matter of fact, this approach locally translates the population share effect that can be modeled in terms of a distribution regression into a quantile effect. Essentially, this method models and then translates population share effects, revealing that the main identifying assumption required is a common trend in *population shares* rather than in *quantiles*. The required assumptions are identical to those recently outlined in Roth and Sant'Anna (2023). The Monte Carlo simulation demonstrates that this approach leads to a substantially reduced bias and RMSE compared to the QTTs estimated using a pooled RIF-DiD approach. Moreover, while the bias is immune to changes in the common trend intensity, the Monte Carlo simulations show that some bias persists for large treatment intensities, as the DDiD-RIF approach still relies on the RIF-inherent first-order approximation.

Secondly, another alternative to the pooled RIF-DiD approach, the QDiD-RIF, is considered. This approach follows the DiD rationale of discretely comparing the quantiles of involved sub-populations to recover a counterfactual quantile for identifying the QTT. Intuitively, this is achieved by updating the RIF for each of the involved sub-populations, an approach previously considered for estimating inequality treatment effects in cross-sectional settings (Rios-Avila and Maroto, 2024). Unlike the latter approach, the required assumption for this approach to be valid is a common trend assumption that holds for quantiles, as outlined in Athey and Imbens (2002, 2006). Naturally, even though overcoming the issue of potentially poor approximations, it is demonstrated that once the assumptions that underlie the QDiD-RIF approach are not accommodated by the DGP, the QTT estimates are substantially biased. The latter is corroborated by the Monte Carlo results, where it is shown that the bias is increasing in the common trend intensity while being immune to large treatment intensities. In summary, while the pooled RIF-DiD approach is structurally misleading, it is shown that there are valid RIF-based alternatives that can be employed to estimate QTTs.

Finally, to examine the performance of the three approaches outside a controlled Monte Carlo simulation environment, we revisit the distributional impact of an expansion in the EITC, as studied in Hoynes and Patel (2018). The comparison of the pooled RIF-DiD approach with the DDiD- and QDiD-RIF approaches reveals that the application of the misleading pooled RIF-DiD approach results in non-negligible underestimations of the QTTs at middle to upper-middle quantiles, where the pooled and marginal distributions differ substantially. Overall, we conclude that the pooled RIF-DiD approach is structurally biased and may lead to significant biases in actual empirical applications, while the two approaches proposed in this paper are able to overcome these structural issues.

Appendix

B.1 Additional formal considerations

B.1.1 Difference between QDiD and DDiD assumptions

In the following section, we outline why a DGP that accommodates the DDiD assumption, as formulated in (3.13) and Roth and Sant'Anna (2023, Proposition 3.2), does not imply that the QTT is identified by approaches relying on a common trend in quantiles, i.e., the QDiD framework, as formulated in Athey and Imbens (2002, 2006). Put differently, we answer the question of why the DGP outlined in the main text (equation (3.26)) prevents the QDiD-RIF approach from producing unbiased QTT estimates, even though approximation errors are mitigated by updating the RIFs for the respective sub-population.

Under the QDiD assumptions outlined in Athey and Imbens (2002, 2006), the potential outcome is generated by the mapping $Y^0(d,t) = h(U,d,t)$, with U referring to unobservables that are unrelated to both D and T. Assume, without loss of generality, that unobservables follow a standard uniform distribution. As noted in Kim and Wooldridge (2024), this is equivalent to a structural assumption that maps unobservables to the potential outcome space by means of the structural function $h(\cdot)$. Using this, the cdf of potential outcomes for the sub-population $(D,T) = (d,t)$ is given by[28]

$$\begin{aligned} F_{Y^0\,|\,dt}(y) &= \mathbb{P}[h(U,d,t) \leq y] \\ &= \int \mathbb{1}\left[h(u,d,t) \leq y\right] f_U(u)\, du. \end{aligned}$$

Assuming that the mapping function is invertible, one can rewrite the argument of the indicator function as $h(u,d,t) \leq y \Leftrightarrow u \leq h^{-1}(y,d,t)$, and hence

$$F_{Y^0\,|\,dt}(y) = \int \mathbb{1}\left[u \leq h^{-1}(y,d,t)\right] f_U(u)\, du.$$

Recalling that U follows a standard uniform distribution, this can be reformulated to show

[28]Note that, since the setting by Athey and Imbens (2002, 2006) implies $U \perp\!\!\!\perp (D,T)$, it holds that $F_{U\,|\,dt}(u) = F_U(u)$.

that the cdf of potential outcomes in terms of the mapping function $h(\cdot)$ can be written as

$$F_{Y_0\,|\,dt}(y) = \int_0^1 \mathbb{1}[u \le h^{-1}(y,d,t)]\; f_U(u)\, du = \int_0^{h^{-1}(y,d,t)} du = h^{-1}(y,d,t),$$

with the second equality arising due to $f_U(u) = 1$ for $u \in [0,1]$ and 0 otherwise. This shows that the entire distribution of potential outcomes can be described by means of $h^{-1}(Y,d,t)$, or equivalently, that the entire quantile function can be described by means of $h(U,d,t)$, i.e., $F^{-1}_{Y^0\,|\,dt}(u) = h(u,d,t)$.

The QDiD assumption essentially revolves around the assumed additivity of $h(U,d,t)$, as described in Athey and Imbens (2002, 2006). Under this assumption, the quantile function can be expressed as:

$$F^{-1}_{Y^0\,|\,dt}(y) = h(u,d,t) = h^T(u,t) + h^D(u,d)$$

which holds for all (d,t) combinations if $U \perp\!\!\!\perp (D,T)$. Hence, assuming that quantiles are additively separable and $U \perp\!\!\!\perp (D,T)$ allows invoking a common trend in quantiles to obtain a counterfactual quantile for the $(D,T) = (1,1)$ sub-population, i.e.,

$$\begin{aligned} h(u,0,1) - h(u,0,0) = h^T(u,1) - h^T(u,0) = h(u,1,1) - h(u,1,0) \\ \implies F^{-1}_{Y^0\,|\,11}(y) = h(u,1,1) = F^{-1}_{Y^0\,|\,10}(y) + \left(F^{-1}_{Y^0\,|\,01}(y) + F^{-1}_{Y^0\,|\,00}(y)\right), \end{aligned} \tag{B.1}$$

where the terms on the RHS of (B.1) are all observable and can, therefore, be estimated.

Based on the QDiD framework, the following reexamines the DDiD framework in terms of the mapping function. Note first that the DDiD assumption as formalized in (3.13) does not imply that $U \perp\!\!\!\perp (D,T)$ but only that $U \perp\!\!\!\perp D\,|\,T = t$ for a fraction θ and $U \perp\!\!\!\perp T\,|\,D = d$ for a fraction $(1-\theta)$. This is another way of stating that the distribution of all possible potential outcomes $(D,T) = (d,t)$ can be represented as a mixture distribution comprising a stationary and an unconfounded sub-population. Using the notion that potential outcomes result from a mapping of unobservables, it is important to recognize that there are essentially two sub-populations for a given (d,t) combination to be considered, a θ as well as a $(1-\theta)$ part. That is, framing the mixture distribution in terms of the mapping function $h(\cdot)$, the different mappings apply for each of the involved sub-population. This incorporates the assumption that $U \perp\!\!\!\perp D\,|\,T = t$ holds for a fraction θ and $U \perp\!\!\!\perp T\,|\,D = d$ holds for a fraction $(1-\theta)$. The potential outcome for the θ part of the mixture distribution thus reads

$$Y^0_\theta(d,t) = h^T(U_\theta,t).$$

Analogously, the $(1-\theta)$ part arises as

$$Y^0_{(1-\theta)}(d,t) = h^D(U_{(1-\theta)},d).$$

Using this, the mixture distribution in (3.13) can be expressed as[29]

$$\begin{aligned}
F_{Y^0\,|\,dt}(y) &= \theta\; G_t(y) + (1-\theta)\; H_d(y) \\
&= \theta \int \mathbb{1}[h^T(u_\theta,t) \leq y]\; dF_{U_\theta\,|\,t}(u) + (1-\theta) \int \mathbb{1}[h^D(u_{(1-\theta)},t) \leq y]\; dF_{U_{(1-\theta)}\,|\,d}(u) \\
&= \theta \int \mathbb{1}[u_\theta \leq \big(h^T(y_\theta,t)\big)^{-1}]\; dF_{U_\theta\,|\,t}(u) \\
&\qquad\qquad + (1-\theta) \int \mathbb{1}[u_{1-\theta} \leq \big(h^D(y_{1-\theta},d)\big)^{-1}]\; dF_{U_{(1-\theta)}\,|\,d}(u)
\end{aligned} \tag{B.2}$$

In principle, one could attempt to follow the QDiD rationale to recover the potential quantile for the treated in the post-treatment period. This is instructive in demonstrating that the DDiD assumption will not permit such recovery. To formally assess this, assume that unobservables follow a standard uniform distribution (similar to above). In this framework, one could consider quantiles of the two isolated fractions of the sub-population:

$$F^{-1}_{Y^0_\theta\,|\,dt}(u) = h^T(u_\theta,t), \quad F^{-1}_{Y^0_{(1-\theta)}\,|\,dt}(u) = h^D(u_{1-\theta},d). \tag{B.3}$$

While this allows for the recovery of the quantiles of both the θ and $(1-\theta)$ parts of a given (d,t) sub-population individually, it does *not* imply that the quantile for the entire (d,t) sub-population is obtainable. This is because the law of total expectations is not applicable to quantiles.

To assess this more formally, start by noting how the setup (B.2) implies a common trend

[29]Note that under the DDiD assumption involving that $U \perp\!\!\!\perp D\,|\,T=t$ for the θ part and $U \perp\!\!\!\perp T\,|\,D=d$ for the $(1-\theta)$ part, this implies that $f_{U\,|\,dt}(u) = f_{U\,|\,T}(u)$ for the θ part, and $f_{U\,|\,dt}(u) = f_{U\,|\,D}(u)$ for the $(1-\theta)$ part.

in potential population shares:

$$
\begin{aligned}
F_{Y^0\,|\,01}(y) - F_{Y^0\,|\,00}(y) = \theta \bigg[&\int \mathbb{1}\big[u_\theta \leq \big(h^T(y_\theta, 1)\big)^{-1}\big] \, dF_{U_\theta\,|\,T=1}(u) \\
&- \int \mathbb{1}\big[u_\theta \leq \big(h^T(y_\theta, 0)\big)^{-1}\big] \, dF_{U_\theta\,|\,T=0}(u) \bigg] \\
= F_{Y^0\,|\,11}(y) - F_{Y^0\,|\,10}(y).
\end{aligned}
$$

Under the DDiD setup, one could now consider the temporal difference in potential outcomes for both groups and examine whether the DDiD setting implies a common trend in quantiles as well:

$$
\begin{aligned}
&F^{-1}_{Y^0\,|\,01}(u) - F^{-1}_{Y^0\,|\,00}(u) \\
&= \bigg[\theta \int \mathbb{1}\big[u_\theta \leq \big(h^T(y_\theta, 1)\big)^{-1}\big] \, dF_{U_\theta\,|\,T=1}(u) \\
&\qquad + (1-\theta) \int \mathbb{1}\big[u_{1-\theta} \leq \big(h^D(y_{1-\theta}, 0)\big)^{-1}\big] \, dF_{U_{1-\theta}\,|\,D=0}(u)\bigg]^{-1} \\
&\quad - \bigg[\theta \int \mathbb{1}\big[u_\theta \leq \big(h^T(y_\theta, 0)\big)^{-1}\big] \, dF_{U_\theta\,|\,T=0}(u) \\
&\qquad + (1-\theta) \int \mathbb{1}\big[u_{1-\theta} \leq \big(h^D(y_{1-\theta}, 0)\big)^{-1}\big] \, dF_{U_{1-\theta}\,|\,D=0}(u)\bigg]^{-1} \\
&\neq F^{-1}_{Y^0\,|\,11}(u) - F^{-1}_{Y^0\,|\,10}(u).
\end{aligned}
$$

The last inequality arises since the group indicator does not easily drop out as in the population share consideration, given that the inverse of the entire term is unequal to the inverse applied individually to the two summands. In other words, the law of total expectations does not apply in terms of quantiles.

To make sense of the finding that for small common trend intensities, approaches that rely on a common trend in quantiles (i.e., the QDiD-RIF approach considered in the main text), it is insightful to consider the cases in which $\theta = 0$ or $\theta = 1$. In these cases, the DGP would be characterized by stationarity or unconfoundedness for the entire population.

Consider the case for $\theta = 1$ (i.e., unconfoundedness):

$$\begin{aligned}
&F^{-1}_{Y^0\,|\,01}(u) - F^{-1}_{Y^0\,|\,00}(u) \\
&\quad = \left[\int \mathbb{1}\left[h^T(u_\theta, 1) \leq y\right] dF_{U_\theta\,|\,T=1}(u)\right]^{-1} - \left[\int \mathbb{1}\left[h^T(u_\theta, 0) \leq y\right] dF_{U_\theta\,|\,T=0}(u)\right]^{-1} \\
&\quad = F^{-1}_{Y^0\,|\,11}(u) - F^{-1}_{Y^0\,|\,10}(u),
\end{aligned}$$

which arises from (B.3) for $\theta = 1$. All this holds vice versa for the case of $\theta = 0$. In sum, if shifts over time and/or across groups are very small – i.e., if the case is somewhat close to either of the border cases – then the performance of approaches that rely on the QDiD assumptions still performs reasonably well. These border cases, in fact, are supported by a QDiD assumption which requires random assignment to a given (d,t) combination, for which the considered cases of population-wide unconfoundedness or stationarity are special cases.

B.1.2 Recovering the QTE using a DDiD-RIF approach

In principle, it is possible to use the two-step approach outlined for the DDiD-RIF case above to approximate effects around the τth quantile of the post-treatment period, $q_{\tau\,|\,1}$, instead of $q_{\tau\,|\,11}$. Put differently, the new target parameter of interest becomes the quantile treatment effect in the post-treatment period, QTE_1, instead of the QTT for $(D,T) = (1,1)$. To estimate the QTE_1, one can adjust the two-step procedure for the QTT outlined in the main text as follows:

(i) Estimate the distributional treatment effect on the treated ('vertical' treatment effect in figure 3.2). Employing a linear DiD regression formulation, this requires to find suitable parameter estimates for the coefficients involved in the following LPM:

$$\mathbb{P}[Y_i \leq q_{\tau\,|\,1}\,|\,D_i, T_i] = \alpha + \beta\, D_i + \delta\, T_i + \gamma\,(D_i \times T_i).$$

Crucially, the threshold in this LPM is $q_{\tau\,|\,1}$ instead of $q_{\tau\,|\,11}$. Using this model, the factual population share in $T = 1$ arises as

$$\mathbb{P}[Y_i \leq q_{\tau\,|\,1}\,|\,D_i, T_i = 1] = \alpha + \beta\, D_i + \delta + \gamma\, D_i,$$

implying that

$$\mathbb{P}[Y_i \leq q_{\tau\,|\,1}\,|\,T_i = 1] = \mathbb{E}_D[\mathbb{P}[Y_i \leq q_{\tau\,|\,1}\,|\,D_i, T_i = 1]] = \alpha + \beta\ \mathbb{E}[D_i] + \delta + \gamma\ \mathbb{E}[D_i].$$

(i.1) Using the latter, the treatment effect on the entire population in $T = 1$ reads

$$\mathbb{P}[Y_i \leq q_{\tau\,|\,1}\,|\,T_i = 1] - \mathbb{P}[Y_i^0 \leq q_{\tau\,|\,1}\,|\,T_i = 1] = \gamma\ \mathbb{E}[D_i].$$

(ii) This aggregated treatment effect in $T = 1$ can be translated into a quantile effect using the RIF rationale, i.e., dividing it by the slope of the cdf of the overall post-treatment cdf, $F_{Y\,|\,1}(y)$, as in (3.19):

$$QTE_1 = -\frac{\gamma\,\mathbb{E}[D_i]}{f_{Y_i\,|\,1}(q_{\tau\,|\,1})}.$$

As in (3.20), the QTE_1 can be estimated by obtaining suitable estimates for γ, $f_{Y_i\,|\,1}(q_{\tau\,|\,1})$, and $\mathbb{E}[D_i]$. Similarly to the remarks in the main text regarding the QTT approximation, it is straightforward to demonstrate that by using the correctly adjusted terms for the RIF approximation, it is possible to find a valid representation of the QTE_1 and the implied counterfactual quantile as in the case of the QTT from (3.18). By using $q_{\tau\,|\,1}$ as the new quantile around which the effect is approximated, this procedure yields a valid representation of both the factual and the counterfactual τth quantile of the entire post-treatment population.

B.2 Conditional RIF-DiD – A reweighting approach

The outline in the main text focuses on the unconditional DiD case to concentrate on the RIF-inherent components that jeopardize a valid approximation of the QTT. Of course, in many applications, the common trend assumption can only be assumed to hold conditionally. In the following, the necessary assumptions for the DTT, as defined in (3.10), to hold conditionally on a vector of covariates $X = x$ is detailed, as outlined in Kim and Wooldridge (2024). Moreover, a suitable reweighting technique is described that can be utilized in the DDiD- and QDiD-RIF approaches. Against the backdrop that RIF approaches have been combined with reweighting techniques in prior contributions in the context of detailed distributional decomposition techniques (Firpo et al., 2018), employing a reweighting technique is straightforward when studying RIF-based DiD approaches.

In the following, the formal results described in Kim and Wooldridge (2024) are essentially paraphrased, replicated, and at points reformulated to suit the RIF approaches outlined in the main text.

Potential distributions under a conditional common trend assumption.

The proposed approach by Kim and Wooldridge (2024) essentially extends the reweighting idea formulated in Abadie (2005). The approach revolves around the notion of aligning the involved sub-populations in terms of observable covariates. To start with, consider the *conditional* version of (3.12):

$$F_{Y_i^0\,|\,x,11}(y) - F_{Y_i^0\,|\,x,10}(y) = F_{Y_i^0\,|\,x,01}(y) - F_{Y_i^0\,|\,x,00}(y),\; y \in \mathscr{Y} \subset \mathbb{R},\; x \in \mathscr{X} \subset \mathbb{R}^K. \quad \text{(B.4)}$$

Note that this conditional formulation allows for non-parallel potential outcome dynamics of different sub-populations in terms of $X = x$.

The second assumption required for identification is an overlapping support assumption, i.e., there must exist a $c > 0$ such that

$$c < \mathbb{P}[D_i = d, T_i = t \,|\, X_i = x] < 1 - c,\; \forall x \in \mathscr{X} \subset \mathbb{R}^K, \quad \text{(B.5)}$$

ensuring that within each of the four sub-populations involved, there is a non-zero probability of finding observationally equivalent observations. Ultimately, this enables the reweighting of individuals. In the following, to simplify the notation, shorthand notation for the involved probability terms is introduced:

$$p_{dt}(x) \equiv \mathbb{P}[D_i = d, T_i = t \,|\, X_i = x], \quad p_{dt} \equiv \mathbb{P}[D_i = d, T_i = t].$$

Given that four sub-populations are involved, it is necessary to be explicit about the sampling distribution. Regarding the sampling procedure, the conditional DiD framework formulated below requires repeated cross-sectional sampling (cf., Kim and Wooldridge, 2024). Specifically, conditional on $T_i = t$, with $t \in \{0,1\}$, draws from the joint distribution (Y_i, D_i, X_i) are assumed to be iid. Ultimately, individuals are assumed to be drawn from the following mixture distribution:

$$P_M(Y_i \leq y, X_i \leq x, D_i = d, T_i = t) = \lambda \cdot t \cdot \mathbb{P}[Y_i \leq y, X_i \leq x, D_i = d \mid T_i = 1] \\ + (1-\lambda)(1-t)\,\mathbb{P}[Y_i \leq y, X_i \leq x, D_i = d \mid T_i = 0], \quad \text{(B.6)}$$

with $\lambda \equiv \mathbb{P}[T_i = 1]$.

As shown in Kim and Wooldridge (2024), given (B.4), (B.5), and (B.6), the counterfactual distribution of the treated in the absence of treatment can be recovered in terms of observable, and hence estimable, elements as follows (extensive proof below):

$$F_{Y_i^0 \mid 11}(y) = \mathbb{E}\left[\frac{p_{11}(x)}{p_{11}}\left(F_{Y_i \mid x,10}(y) + F_{Y_i \mid x,01}(y) - F_{Y_i \mid x,00}(y)\right)\right]. \quad \text{(B.7)}$$

This result and the implied idea of reweighting sub-populations (see below) can be utilized in a RIF application, similar to the distributional decomposition approaches formulated in Firpo and Pinto (2016) and Firpo et al. (2018). Like a hybrid RIF decomposition, the proposed approach involves the isolated consideration of sub-populations and the corresponding reweighting of these groups. Establishing this correspondence highlights that the proposed reweighting approach in the context of the DDiD- and QDiD-RIF approaches is readily implementable using well-established reweighting techniques.

In the following, to facilitate readability, the simplified notation of Kim and Wooldridge (2024) is used, i.e., define $D_{i1} \equiv (1-D_i)(1-T_i)$, $D_{i2} \equiv (1-D_i)T_i$, $D_{i3} \equiv D_i(1-T_i)$, and $D_{i0} \equiv D_i T_i$. Further, denote the conditional and unconditional propensities of belonging to group j as $p_j(x) \equiv \mathbb{P}[D_{ij} = 1 \mid X_i = x]$ and $p_j \equiv \mathbb{P}[D_{ij} = 1]$, respectively. Furthermore, denote the reweighted sub-population from group D_{ij} to the treated group in the post-treatment period, D_{i0}, as $F_{Y_i^{11} \mid D_j}(y)$. Using this notation, as noted in Kim and Wooldridge (2024) and outlined explicitly below, the sub-populations in (B.7) can be rewritten as follows:

$$F_{Y_i^{11}\,|\,D_j}(y) = \int_{i\in j} F_{Y_i\,|\,D_j,x}\ \omega_j(x)\ dF_{X_i\,|\,D_j}, \quad \text{with} \quad \omega_j(x) \equiv \frac{p_0(x)}{p_j(x)}\,\frac{p_j}{p_0}. \tag{B.8}$$

Crucially, the resulting reweighting factors, $\omega_j(x)$, are structurally identical to those derived in DiNardo et al. (1996). Intuitively, the reweighting factors assign more weight to observations from the respective group j that are similar, in terms of their covariate endowment, to observations in group 0 (large $p_0(x)$) and, at the same time, have a small probability of being observed in group j (small $p_j(x)$). The weights can be estimated using, for example, a multinomial logit model with D_{i1} as the base group.[30] In summary, the advantage of this approach is its ability to align agents within a separate reweighting exercise, ensuring that agents are observationally similar.

Counterfactual distribution under a common trend assumption

The following is an extensive replication of the proof for identifying the counterfactual distribution as in (B.7) suggested by Kim and Wooldridge (2024).

Proof – Identification of counterfactual distribution in post-treatment period. In the following, all population quantities are with respect to the mixture distribution defined in (B.6). Start by writing the unconditional distribution of the counterfactual potential outcome of treated individuals in the post-treatment period if they had not been treated as

$$\begin{aligned} F_{Y_i^0\,|\,11}(y) &= \mathbb{P}[Y_i^0 \leq y\,|\,D_i = 1, T_i = 1] = \mathbb{P}[Y_i^0 \leq y, D_i = 1, T_i = 1] \cdot \frac{1}{\mathbb{P}[D_i = 1, T_i = 1]} \\ &= \frac{1}{p_{11}} \cdot \mathbb{E}\left[\mathbb{P}[Y_i^0 \leq y, D_i = 1, T_i = 1\,|\,X_i = x]\right]. \end{aligned} \tag{B.9}$$

The inner probability can further be written as

$$\mathbb{P}[Y_i^0 \leq y, D_i = 1, T_i = 1\,|\,X_i = x] = \mathbb{E}[D_i \cdot T_i \cdot \mathbb{1}[Y_i^0 \leq y]\,|\,X_i = x].$$

Using this notation and using $\mathbb{P}[A \cap B\,|\,X_i = x] = \mathbb{P}[A\,|\,B, X_i = x] \cdot \mathbb{P}[B\,|\,X_i = x]$ one can

[30]This is described in detail in Kim and Wooldridge (2024). See also below.

write $\mathbb{P}[Y_i^0 \leq y, D_i = 1, T_i = 1 \,|\, X_i = x]$ as

$$\mathbb{E}[D_i \cdot T_i \cdot \mathbb{1}[Y_i^0 \leq y] \,|\, X_i = x] = \mathbb{E}[D_i \cdot T_i \cdot \mathbb{1}[Y_i^0 \leq y] \,|\, X_i = x, D_i = 1, T_i = 1] \cdot \mathbb{P}[D_i = 1, T_i = 1 \,|\, X_i = x],$$

or written in terms of probability operators:

$$\begin{aligned}\mathbb{P}[Y_i^0 \leq y, D_i = 1, T_i = 1 \,|\, X_i = x] &= p_{11}(x) \cdot \mathbb{P}[Y_i^0 \leq y \,|\, X_i = x, D_i = 1, T_i = 1] \\ &= p_{11}(x) \cdot F_{Y^0 \,|\, x,11}(y).\end{aligned}$$

Inserting this result into (B.9) yields

$$F_{Y_i^0 \,|\, 11}(y) = \mathbb{E}\left[\frac{p_{11}(x)}{p_{11}} \cdot F_{Y_i^0 \,|\, x,11}(y)\right]. \tag{B.10}$$

Given (B.5), the conditional propensity $p_{11}(x)$ is well defined. Further, under the conditional common trend assumption from (B.4), it follows that the RHS of (B.10) can be written in terms of observable elements, i.e.

$$F_{Y_i^0 \,|\, 11}(y) = \mathbb{E}\left\{\frac{p_{11}(x)}{p_{11}} \underbrace{\left[F_{Y_i \,|\, x,10}(y) + F_{Y_i \,|\, x,01}(y) - F_{Y_i \,|\, x,00}(y)\right]}_{=F_{Y_i^0 \,|\, x,11}(y), \text{ by (B.4)}}\right\},$$

which completes the proof. □

Reweighting by sub-populations

As described below, using the reweighting factors in the context of the DDiD- and QDiD-RIF approaches can be achieved by formulating sub-population-specific reweighting factors, $\omega_j(x)$, and running the unconditional analyses on the reweighted sample. It is therefore particularly convenient to derive these sub-population-specific reweighting factors with which the entire sample can be adjusted (compare procedures in DiNardo et al., 1996; Firpo et al., 2018). To reweight all individuals from the involved sub-populations separately, note that the sampling distribution, as in (B.6), allows for changing compositions of groups over time. Thus, each of the four sub-populations needs to be reweighted such that the covariate endowment is similar to that of the treated observations in the post-treatment period.[31]

To derive the sub-population specific reweighting factors, start by noting that each of the involved conditional distributions in (B.7) can be written as

$$
\begin{aligned}
F_{Y_i\,|\,x,D_{ij}=1}(y) &= \mathbb{E}[\mathbb{1}[Y_i \leq y]\,|\,X_i = x, D_{ij}=1] = \mathbb{E}[D_{ij}\cdot \mathbb{1}[Y_i \leq y]\,|\,X_i = x, D_{ij}=1] \\
&= \frac{\mathbb{E}[D_{ij}\cdot \mathbb{1}[Y_i \leq y\,|\,X_i = x]}{p_j(x)},
\end{aligned}
$$

where the second line is a consequence of $\mathbb{E}[Z\cdot D_j\,|\,X_i = x] = p_j(x)\cdot\mathbb{E}[Z\cdot D_j\,|\,D_j = 1, X_i = x] + (1-p_j(x))\cdot\mathbb{E}[Z\cdot D_j\,|\,D_j = 0, X_i = x]$, and since $\mathbb{E}[Z\cdot D_j\,|\,D_j = 0, X_i = x] = 0$ it follows that $(1/p_j(x))\cdot\mathbb{E}[Z\cdot D_j\,|\,X_i = x] = \mathbb{E}[Z\cdot D_j\,|\,D_j = 1, X_i = x]$, with Z being some random variable.

Using this, the counterfactual distribution from (B.7) can be expressed as

$$
\begin{aligned}
&F_{Y_i^0\,|\,11}(y) = \\
&\quad\mathbb{E}\left\{\frac{p_0(x)}{p_0}\left(\frac{\mathbb{E}[D_{i3}\cdot\mathbb{1}[Y_i\leq y]\,|\,X_i = x]}{p_3(x)} + \frac{\mathbb{E}[D_{i2}\cdot\mathbb{1}[Y_i\leq y]\,|\,X_i = x]}{p_2(x)}\right.\right. \\
&\qquad\qquad\qquad\qquad\left.\left. - \frac{\mathbb{E}[D_{i1}\cdot\mathbb{1}[Y_i\leq y]\,|\,X_i = x]}{p_1(x)}\right)\right\} \\
&= \frac{1}{p_0}\left(\mathbb{E}\left\{\mathbb{E}\left[\frac{p_0(x)}{p_3(x)}\cdot D_{i3}\mathbb{1}[Y_i\leq y]\,|\,X_i = x\right] + \mathbb{E}\left[\frac{p_0(x)}{p_2(x)}\cdot D_{i2}\mathbb{1}[Y_i\leq y]\,|\,X_i = x\right]\right.\right. \\
&\qquad\qquad\qquad\qquad\left.\left. - \mathbb{E}\left[\frac{p_0(x)}{p_1(x)}\cdot D_{i1}\mathbb{1}[Y_i\leq y]\,|\,X_i = x\right]\right\}\right) \\
&= \frac{1}{p_0}\left(\mathbb{E}\left[\frac{p_0(x)}{p_3(x)}\cdot D_{i3}\mathbb{1}[Y_i\leq y]\right] + \mathbb{E}\left[\frac{p_0(x)}{p_2(x)}\cdot D_{i2}\mathbb{1}[Y_i\leq y]\right] - \mathbb{E}\left[\frac{p_0(x)}{p_1(x)}\cdot D_{i1}\mathbb{1}[Y_i\leq y]\right]\right).
\end{aligned}
\tag{B.11}
$$

[31] Of course, the change in the covariate endowment over time must not be related to the treatment assignment, which would essentially make them ‘bad controls’.

To ultimately express the counterfactual cdf in terms of a reweighting exercise that utilizes sub-population-specific weights, reconsider the jth reweighted cdf from (B.11) separately:

$$\begin{aligned}
F_{Y_i^{11}\,|\,D_j}(y) &= \mathbb{E}\left\{ \frac{p_0(x)}{p_0} \frac{1}{p_j(x)} \cdot D_{ij}\mathbb{1}\left[Y_i \le y\right] \right\} \\
&= \mathbb{E}\left\{ \frac{p_0(x)}{p_0} \frac{1}{p_j(x)} \cdot D_{ij}\mathbb{1}\left[Y_i \le y\right] \,|\, D_j \right\} \cdot p_j \\
&= \mathbb{E}\left\{ \mathbb{E}\left[\frac{p_0(x)}{p_0} \frac{p_j}{p_j(x)} \mathbb{1}\left[Y_i \le y\right] \,|\, D_j, X_i = x \right] \right\} \\
&= \mathbb{E}\left\{ \frac{p_0(x)}{p_0} \frac{p_j}{p_j(x)} \mathbb{E}\left[\mathbb{1}\left[Y_i \le y\right] \,|\, D_j, X_i = x \right] \right\} \\
&= \mathbb{E}\left\{ F_{Y\,|\,D_j,x}\ \omega_j(x) \right\},
\end{aligned}$$

$$\Longrightarrow F_{Y_i^{11}\,|\,D_j}(y) = \int_{i\in j} F_{Y_i\,|\,D_j,x}\ \omega_j(x)\ dF_{X_i\,|\,D_j}, \quad \text{with} \quad \omega_j(x) \equiv \frac{p_0(x)}{p_j(x)} \frac{p_j}{p_0}, \tag{B.12}$$

where the second line is a consequence of $\mathbb{E}[Z \cdot D_j] = p_j \cdot \mathbb{E}[Z \cdot D_j \,|\, D_j = 1] + (1 - p_j) \cdot \mathbb{E}[Z \cdot D_j \,|\, D_j = 0]$, and since $\mathbb{E}[Z \cdot D_j \,|\, D_j = 0] = 0$ it follows that $\mathbb{E}[Z \cdot D_j] = p_j \cdot \mathbb{E}[Z \cdot D_j \,|\, D_j = 1]$, with Z being some random variable. This makes it apparent that, given an appropriate model for the involved conditional cdf, $F_{Y\,|\,D_j,x}$, the unconditional reweighted cdf could, for example, be obtained by means of distribution regression techniques involving the corresponding weights $\omega_j(x)$.[32] Regarding the estimation of the reweights, discrete choice models can be used (DiNardo et al., 1996; Cattaneo, 2010; Kim and Wooldridge, 2024, see below for a detailed description). This notation further facilitates the expression of reweighted distributional statistics using the corresponding RIF:

$$\nu(F_{Y_i^{11}} \,|\, D_j) = \int_{i\in j} \mathbb{E}[RIF(Y_i; \nu, F_{Y_i^{11}\,|\,D_j}(y)) \,|\, X_i = x]\ \omega_j(x)\ dF_{X_i\,|\,D_j}. \tag{B.13}$$

This shows that the distributional statistic $\nu(\cdot)$ can be expressed by means of a simple weighted average using only observations from the respective sub-population.

Estimation of involved elements

[32] For example, in the case of a linear model for $F_{Y\,|\,D_j,x}$ one would use a weighted least squares approach.

The reweighting factors can be easily estimated using multinomial logistic regression techniques (DiNardo et al., 1996; Cattaneo, 2010; Kim and Wooldridge, 2024). Concretely, the weighting factors arise as

$$\widehat{\omega}_j(x) = \begin{cases} \frac{1}{\widehat{p_0}} & \text{if } j = 0 \\ \frac{\widehat{p_0}(x)}{\widehat{p_1}(x)} \cdot \frac{1}{\widehat{p_0}} & \text{if } j = 1 \\ \frac{\widehat{p_0}(x)}{\widehat{p_2}(x)} \cdot \frac{1}{\widehat{p_0}} & \text{if } j = 2 \\ \frac{\widehat{p_0}(x)}{\widehat{p_3}(x)} \cdot \frac{1}{\widehat{p_0}} & \text{if } j = 3, \end{cases}$$

with the involved propensities being estimable via

$$\log\left(\frac{p_j(X_i)}{p_0(X_i)}\right) = R_K(X_i)'\gamma_{K,j} + error_i,$$

with $R_K(X_i)$ being a vector of K basis terms such as polynomials (cf., Kim and Wooldridge, 2024).

RIF-DiD and reweighting

In the case of the DDiD-RIF approach discussed above, the underlying reweighting idea corresponds to the one formulated in (B.11). While the formulation in (B.11) implies a non-parametric construction of the cdf, estimating parameters from the saturated regression formulation, as implied by (3.20), using weighted least squares (WLS) with $\widehat{\omega}_j(x)$ as weights in the implied first-stage distribution regression, is a viable option. Note that the kernel density in the denominator of (3.20) does not need to be adjusted since it exclusively concerns the distribution of the treated in the post-treatment period, which does not require reweighting.

Regarding the corresponding reweighting scheme for the QDiD-RIF approach, note that (B.13) represents the suitably reweighted unconditional distributional statistic, $v(F_{Y_i^{11}} \mid D_j)$. Hence, using the reweighted compound RIFs as in (3.21) renders the reweighting approach feasible and easily implementable in the context of the QDiD-RIF approach as well.

B.3 Additional results

B.3.1 Monte Carlo simulation – full results

— Intentionally left blank —

Table B.1: Results for all three approaches. Common trend intensity $c = 0.1$, treatment intensity $r = 2.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.263	0.236	0.359	0.272	0.179	0.285	0.223	0.239	0.358	0.270	0.287	0.432	0.317	-0.028	0.096	0.008	0.041	0.058	0.036	0.050	0.112	0.037
15	0.406	0.348	0.510	0.418	0.288	0.429	0.361	0.351	0.511	0.420	0.404	0.592	0.471	-0.058	0.104	0.013	0.046	0.063	0.043	0.074	0.121	0.045
20	0.557	0.463	0.667	0.574	0.400	0.580	0.514	0.464	0.668	0.575	0.525	0.750	0.635	-0.094	0.111	0.017	0.049	0.068	0.048	0.106	0.130	0.051
25	0.716	0.591	0.834	0.737	0.518	0.727	0.662	0.593	0.834	0.739	0.661	0.938	0.811	-0.125	0.118	0.021	0.056	0.079	0.057	0.137	0.142	0.061
30	0.885	0.731	1.008	0.908	0.648	0.891	0.817	0.734	1.009	0.910	0.816	1.124	0.996	-0.154	0.123	0.023	0.066	0.088	0.067	0.167	0.151	0.071
35	1.064	0.889	1.190	1.085	0.795	1.064	0.982	0.890	1.186	1.088	0.981	1.312	1.180	-0.175	0.126	0.022	0.072	0.095	0.076	0.189	0.158	0.079
40	1.253	1.061	1.381	1.275	0.960	1.243	1.170	1.063	1.378	1.274	1.157	1.518	1.379	-0.192	0.128	0.021	0.080	0.107	0.083	0.208	0.167	0.086
45	1.454	1.252	1.577	1.475	1.126	1.420	1.345	1.250	1.576	1.474	1.377	1.730	1.600	-0.202	0.122	0.020	0.092	0.121	0.096	0.222	0.172	0.098
50	1.667	1.458	1.780	1.680	1.319	1.600	1.532	1.460	1.782	1.680	1.587	1.969	1.825	-0.209	0.113	0.013	0.105	0.141	0.112	0.234	0.181	0.113
55	1.892	1.681	1.996	1.905	1.511	1.792	1.737	1.690	2.000	1.909	1.843	2.198	2.083	-0.211	0.105	0.014	0.125	0.160	0.133	0.246	0.191	0.133
60	2.130	1.928	2.214	2.138	1.736	1.988	1.938	1.933	2.225	2.142	2.118	2.430	2.347	-0.202	0.084	0.008	0.146	0.171	0.153	0.249	0.191	0.153
65	2.382	2.189	2.446	2.385	1.967	2.170	2.168	2.189	2.457	2.385	2.394	2.714	2.618	-0.193	0.064	0.003	0.161	0.201	0.173	0.251	0.211	0.173
70	2.648	2.484	2.692	2.643	2.211	2.365	2.372	2.486	2.703	2.654	2.747	2.986	2.907	-0.165	0.044	-0.006	0.198	0.244	0.209	0.257	0.247	0.209
75	2.930	2.805	2.938	2.913	2.464	2.554	2.560	2.814	2.947	2.929	3.139	3.305	3.250	-0.125	0.008	-0.017	0.254	0.289	0.265	0.282	0.289	0.265
80	3.227	3.148	3.207	3.192	2.729	2.716	2.744	3.155	3.216	3.221	3.559	3.662	3.602	-0.079	-0.020	-0.034	0.319	0.362	0.332	0.329	0.363	0.333
85	3.540	3.524	3.465	3.485	2.950	2.852	2.882	3.529	3.462	3.496	4.053	4.053	4.043	-0.016	-0.075	-0.055	0.425	0.464	0.448	0.425	0.470	0.451
90	3.870	3.902	3.763	3.764	3.050	2.944	2.967	3.925	3.723	3.732	4.707	4.627	4.626	0.032	-0.107	-0.106	0.626	0.663	0.643	0.626	0.671	0.651

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.2: Results for all three approaches. Common trend intensity $c = 0.1$, treatment intensity $r = 7.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.790	0.609	1.291	0.798	0.556	1.191	0.750	0.610	1.295	0.797	0.659	1.380	0.844	-0.181	0.501	0.008	0.041	0.073	0.036	0.185	0.506	0.037
15	1.218	0.911	1.802	1.230	0.856	1.701	1.173	0.912	1.804	1.232	0.969	1.904	1.283	-0.306	0.584	0.013	0.046	0.080	0.043	0.310	0.589	0.045
20	1.670	1.201	2.326	1.687	1.138	2.204	1.628	1.203	2.330	1.688	1.262	2.440	1.748	-0.469	0.656	0.017	0.049	0.089	0.048	0.472	0.662	0.051
25	2.148	1.563	2.859	2.169	1.490	2.727	2.094	1.563	2.863	2.171	1.631	2.982	2.243	-0.586	0.711	0.021	0.056	0.098	0.057	0.588	0.717	0.061
30	2.655	1.958	3.400	2.678	1.873	3.263	2.586	1.962	3.405	2.681	2.038	3.526	2.767	-0.697	0.745	0.023	0.066	0.104	0.067	0.700	0.752	0.071
35	3.192	2.413	3.941	3.213	2.317	3.784	3.109	2.413	3.946	3.216	2.505	4.081	3.307	-0.779	0.749	0.022	0.072	0.115	0.076	0.782	0.758	0.079
40	3.760	2.919	4.480	3.782	2.811	4.308	3.676	2.920	4.485	3.781	3.025	4.640	3.886	-0.841	0.720	0.022	0.081	0.128	0.083	0.845	0.731	0.086
45	4.362	3.484	5.009	4.383	3.357	4.824	4.253	3.490	5.017	4.382	3.601	5.183	4.508	-0.878	0.647	0.021	0.093	0.143	0.096	0.883	0.663	0.098
50	5.000	4.119	5.538	5.013	3.971	5.334	4.866	4.121	5.540	5.014	4.260	5.737	5.158	-0.881	0.538	0.013	0.108	0.155	0.112	0.888	0.560	0.113
55	5.675	4.818	6.061	5.689	4.653	5.833	5.521	4.825	6.059	5.692	4.975	6.283	5.866	-0.857	0.386	0.014	0.126	0.168	0.133	0.866	0.421	0.133
60	6.390	5.602	6.582	6.398	5.411	6.318	6.198	5.604	6.575	6.402	5.789	6.827	6.607	-0.788	0.192	0.008	0.143	0.193	0.153	0.801	0.272	0.153
65	7.146	6.462	7.098	7.149	6.247	6.799	6.932	6.461	7.092	7.149	6.670	7.393	7.382	-0.684	-0.048	0.003	0.163	0.222	0.173	0.703	0.227	0.173
70	7.945	7.425	7.605	7.939	7.144	7.267	7.669	7.430	7.614	7.951	7.684	7.917	8.204	-0.520	-0.340	-0.006	0.203	0.251	0.209	0.559	0.422	0.209
75	8.789	8.479	8.124	8.772	8.133	7.728	8.419	8.480	8.128	8.789	8.804	8.499	9.109	-0.310	-0.665	-0.017	0.255	0.292	0.265	0.402	0.726	0.265
80	9.680	9.615	8.645	9.646	9.193	8.169	9.198	9.627	8.652	9.675	10.013	9.087	10.057	-0.065	-1.035	-0.034	0.325	0.354	0.332	0.331	1.094	0.333
85	10.620	10.863	9.211	10.565	10.329	8.642	9.962	10.880	9.204	10.576	11.401	9.779	11.123	0.243	-1.409	-0.055	0.424	0.457	0.448	0.489	1.481	0.451
90	11.610	12.175	9.892	11.504	11.322	9.144	10.707	12.165	9.868	11.472	13.003	10.709	12.366	0.565	-1.718	-0.106	0.633	0.627	0.643	0.848	1.829	0.651

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.3: Results for all three approaches. Common trend intensity $c = 0.25$, treatment intensity $r = 2.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.263	0.180	0.343	0.347	0.122	0.258	0.292	0.182	0.344	0.350	0.233	0.426	0.395	-0.083	0.079	0.084	0.042	0.067	0.039	0.093	0.104	0.092
15	0.406	0.281	0.491	0.519	0.221	0.401	0.462	0.283	0.490	0.519	0.347	0.585	0.579	-0.125	0.085	0.113	0.049	0.071	0.046	0.134	0.111	0.122
20	0.557	0.381	0.640	0.698	0.313	0.544	0.635	0.382	0.637	0.697	0.446	0.740	0.767	-0.176	0.084	0.141	0.052	0.078	0.052	0.184	0.115	0.151
25	0.716	0.502	0.793	0.881	0.426	0.683	0.799	0.499	0.794	0.883	0.575	0.909	0.965	-0.214	0.077	0.165	0.060	0.088	0.063	0.223	0.116	0.177
30	0.885	0.632	0.961	1.070	0.542	0.834	0.975	0.633	0.959	1.073	0.719	1.082	1.166	-0.253	0.076	0.185	0.068	0.094	0.073	0.262	0.121	0.198
35	1.064	0.781	1.137	1.261	0.680	1.002	1.151	0.781	1.141	1.264	0.878	1.270	1.363	-0.283	0.073	0.197	0.076	0.105	0.083	0.293	0.128	0.214
40	1.253	0.948	1.322	1.459	0.836	1.162	1.345	0.947	1.321	1.457	1.056	1.464	1.574	-0.305	0.068	0.206	0.086	0.118	0.091	0.317	0.136	0.225
45	1.454	1.130	1.511	1.662	1.005	1.334	1.522	1.130	1.512	1.661	1.252	1.672	1.795	-0.324	0.057	0.208	0.100	0.131	0.106	0.339	0.143	0.233
50	1.667	1.335	1.708	1.863	1.182	1.504	1.702	1.338	1.715	1.866	1.475	1.902	2.020	-0.332	0.041	0.197	0.114	0.154	0.123	0.351	0.159	0.232
55	1.892	1.559	1.917	2.077	1.377	1.687	1.894	1.562	1.921	2.084	1.724	2.126	2.257	-0.333	0.025	0.186	0.133	0.163	0.144	0.359	0.165	0.235
60	2.130	1.811	2.131	2.290	1.601	1.900	2.059	1.820	2.130	2.288	2.003	2.377	2.515	-0.319	0.001	0.160	0.155	0.182	0.168	0.354	0.182	0.232
65	2.382	2.084	2.367	2.508	1.846	2.085	2.259	2.086	2.368	2.511	2.300	2.660	2.764	-0.298	-0.015	0.126	0.176	0.221	0.191	0.346	0.221	0.229
70	2.648	2.406	2.603	2.723	2.129	2.253	2.416	2.402	2.610	2.734	2.687	2.918	3.024	-0.243	-0.045	0.075	0.213	0.261	0.231	0.323	0.264	0.243
75	2.930	2.757	2.847	2.932	2.385	2.431	2.552	2.772	2.844	2.946	3.090	3.247	3.272	-0.173	-0.083	0.003	0.269	0.310	0.290	0.319	0.320	0.290
80	3.227	3.140	3.115	3.132	2.681	2.623	2.660	3.144	3.125	3.144	3.594	3.598	3.593	-0.087	-0.112	-0.095	0.342	0.382	0.364	0.353	0.398	0.375
85	3.540	3.574	3.353	3.298	3.004	2.702	2.644	3.592	3.354	3.288	4.120	4.005	3.929	0.034	-0.187	-0.242	0.446	0.497	0.494	0.447	0.531	0.550
90	3.870	4.049	3.677	3.404	3.169	2.773	2.478	4.046	3.663	3.410	4.878	4.590	4.310	0.179	-0.193	-0.466	0.659	0.709	0.710	0.682	0.734	0.848

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.4: Results for all three approaches. Common trend intensity $c = 0.25$, treatment intensity $r = 7.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.790	0.488	1.229	0.874	0.430	1.122	0.819	0.490	1.233	0.876	0.543	1.332	0.922	-0.302	0.439	0.084	0.043	0.082	0.039	0.305	0.446	0.093
15	1.218	0.752	1.724	1.332	0.691	1.616	1.277	0.752	1.726	1.332	0.814	1.831	1.389	-0.466	0.507	0.114	0.048	0.088	0.046	0.468	0.514	0.123
20	1.670	1.024	2.218	1.812	0.955	2.100	1.752	1.023	2.218	1.812	1.089	2.338	1.880	-0.646	0.548	0.142	0.053	0.095	0.052	0.649	0.556	0.151
25	2.148	1.354	2.740	2.314	1.277	2.599	2.238	1.354	2.745	2.314	1.423	2.873	2.395	-0.794	0.592	0.165	0.058	0.106	0.062	0.796	0.601	0.177
30	2.655	1.732	3.273	2.840	1.647	3.118	2.753	1.734	3.276	2.844	1.823	3.413	2.930	-0.923	0.618	0.185	0.068	0.116	0.071	0.925	0.629	0.199
35	3.192	2.165	3.811	3.391	2.061	3.640	3.276	2.166	3.810	3.393	2.263	3.967	3.495	-1.026	0.619	0.200	0.077	0.129	0.083	1.029	0.633	0.216
40	3.760	2.660	4.346	3.970	2.540	4.148	3.850	2.664	4.350	3.970	2.769	4.530	4.088	-1.100	0.586	0.210	0.089	0.141	0.091	1.103	0.603	0.229
45	4.362	3.219	4.882	4.574	3.082	4.673	4.435	3.222	4.882	4.573	3.347	5.082	4.713	-1.144	0.520	0.212	0.104	0.156	0.107	1.148	0.542	0.237
50	5.000	3.850	5.418	5.204	3.691	5.192	5.033	3.854	5.414	5.204	3.993	5.631	5.366	-1.150	0.418	0.204	0.117	0.168	0.125	1.156	0.450	0.239
55	5.675	4.558	5.951	5.868	4.374	5.711	5.678	4.567	5.957	5.872	4.728	6.187	6.046	-1.117	0.276	0.193	0.137	0.183	0.145	1.126	0.331	0.241
60	6.390	5.352	6.496	6.560	5.146	6.226	6.332	5.354	6.492	6.560	5.567	6.760	6.787	-1.038	0.106	0.170	0.158	0.210	0.169	1.050	0.235	0.240
65	7.146	6.252	7.036	7.285	6.020	6.710	7.034	6.248	7.048	7.286	6.489	7.332	7.556	-0.894	-0.110	0.139	0.182	0.239	0.196	0.912	0.262	0.240
70	7.945	7.270	7.575	8.033	6.983	7.184	7.718	7.269	7.594	8.041	7.557	7.938	8.340	-0.675	-0.370	0.088	0.223	0.283	0.237	0.711	0.466	0.253
75	8.789	8.413	8.118	8.808	8.046	7.717	8.418	8.418	8.127	8.820	8.763	8.520	9.166	-0.376	-0.671	0.019	0.274	0.321	0.293	0.465	0.743	0.294
80	9.680	9.684	8.695	9.601	9.192	8.188	9.118	9.697	8.708	9.623	10.154	9.194	10.065	0.004	-0.985	-0.079	0.357	0.394	0.370	0.356	1.061	0.378
85	10.620	11.118	9.293	10.406	10.525	8.609	9.737	11.153	9.320	10.425	11.725	9.917	11.030	0.498	-1.327	-0.213	0.468	0.507	0.501	0.683	1.420	0.544
90	11.610	12.717	9.988	11.185	11.814	9.140	10.247	12.713	9.964	11.208	13.572	10.885	12.109	1.107	-1.622	-0.425	0.683	0.692	0.714	1.301	1.763	0.830

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.5: Results for all three approaches. Common trend intensity $c = 0.5$, treatment intensity $r = 2.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.263	0.120	0.332	0.654	0.059	0.226	0.596	0.118	0.330	0.654	0.182	0.442	0.712	-0.144	0.068	0.391	0.047	0.085	0.043	0.151	0.109	0.393
15	0.406	0.190	0.469	0.932	0.129	0.358	0.871	0.187	0.466	0.932	0.252	0.593	0.999	-0.216	0.063	0.526	0.051	0.092	0.051	0.222	0.111	0.529
20	0.557	0.282	0.611	1.207	0.209	0.489	1.135	0.280	0.608	1.205	0.355	0.743	1.286	-0.275	0.055	0.651	0.059	0.100	0.059	0.281	0.114	0.653
25	0.716	0.385	0.769	1.477	0.305	0.624	1.394	0.378	0.764	1.475	0.477	0.911	1.574	-0.332	0.052	0.761	0.071	0.110	0.071	0.339	0.122	0.764
30	0.885	0.506	0.933	1.739	0.408	0.775	1.634	0.503	0.933	1.737	0.608	1.087	1.850	-0.379	0.048	0.854	0.078	0.119	0.082	0.387	0.128	0.857
35	1.064	0.646	1.110	1.991	0.528	0.931	1.869	0.642	1.110	1.994	0.764	1.276	2.112	-0.418	0.046	0.927	0.090	0.134	0.096	0.428	0.141	0.932
40	1.253	0.806	1.294	2.228	0.672	1.117	2.093	0.805	1.296	2.227	0.941	1.472	2.364	-0.447	0.040	0.975	0.101	0.145	0.108	0.459	0.150	0.981
45	1.454	0.992	1.481	2.444	0.830	1.284	2.287	0.995	1.483	2.440	1.141	1.688	2.601	-0.462	0.027	0.989	0.118	0.160	0.124	0.477	0.162	0.997
50	1.667	1.200	1.681	2.632	1.037	1.467	2.448	1.202	1.687	2.626	1.364	1.908	2.818	-0.466	0.015	0.965	0.133	0.177	0.146	0.485	0.177	0.976
55	1.892	1.438	1.891	2.800	1.233	1.660	2.585	1.440	1.886	2.795	1.637	2.147	3.028	-0.454	-0.001	0.908	0.154	0.193	0.169	0.479	0.193	0.924
60	2.130	1.707	2.121	2.933	1.490	1.851	2.685	1.713	2.116	2.920	1.934	2.408	3.195	-0.423	-0.009	0.803	0.175	0.221	0.193	0.458	0.221	0.826
65	2.382	2.013	2.362	3.026	1.772	2.020	2.733	2.008	2.372	3.012	2.270	2.699	3.330	-0.369	-0.020	0.644	0.200	0.268	0.223	0.420	0.269	0.682
70	2.648	2.376	2.600	3.059	2.062	2.196	2.713	2.375	2.606	3.066	2.694	2.998	3.433	-0.273	-0.049	0.411	0.246	0.314	0.278	0.367	0.317	0.496
75	2.930	2.797	2.860	3.012	2.401	2.391	2.549	2.803	2.862	3.017	3.191	3.327	3.464	-0.133	-0.070	0.083	0.312	0.371	0.350	0.339	0.377	0.359
80	3.227	3.275	3.131	2.874	2.772	2.574	2.272	3.262	3.160	2.884	3.784	3.674	3.421	0.048	-0.096	-0.353	0.399	0.452	0.450	0.402	0.462	0.572
85	3.540	3.844	3.407	2.580	3.151	2.656	1.809	3.864	3.404	2.620	4.501	4.169	3.328	0.304	-0.133	-0.960	0.528	0.597	0.596	0.609	0.611	1.130
90	3.870	4.506	3.716	2.011	3.470	2.583	0.920	4.532	3.689	2.011	5.556	4.863	3.145	0.636	-0.154	-1.859	0.804	0.859	0.855	1.025	0.872	2.045

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.6: Results for all three approaches. Common trend intensity $c = 0.5$, treatment intensity $r = 7.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.790	0.333	1.181	1.182	0.267	1.050	1.126	0.335	1.186	1.183	0.394	1.306	1.237	-0.457	0.391	0.392	0.050	0.103	0.042	0.460	0.405	0.394
15	1.218	0.525	1.668	1.744	0.456	1.526	1.682	0.526	1.667	1.745	0.588	1.817	1.811	-0.693	0.450	0.527	0.052	0.115	0.052	0.695	0.465	0.529
20	1.670	0.781	2.179	2.321	0.703	2.011	2.247	0.781	2.186	2.324	0.855	2.338	2.397	-0.889	0.509	0.651	0.061	0.124	0.059	0.891	0.524	0.654
25	2.148	1.071	2.712	2.909	0.982	2.526	2.822	1.073	2.716	2.912	1.161	2.884	3.002	-1.077	0.563	0.761	0.071	0.135	0.070	1.079	0.579	0.764
30	2.655	1.421	3.263	3.508	1.322	3.061	3.397	1.422	3.270	3.509	1.522	3.447	3.614	-1.234	0.608	0.853	0.079	0.148	0.081	1.236	0.625	0.857
35	3.192	1.833	3.818	4.116	1.707	3.621	3.988	1.837	3.820	4.122	1.954	4.012	4.233	-1.358	0.627	0.924	0.092	0.156	0.094	1.361	0.646	0.929
40	3.760	2.310	4.382	4.731	2.172	4.155	4.594	2.312	4.388	4.733	2.443	4.587	4.862	-1.450	0.622	0.971	0.104	0.170	0.106	1.454	0.644	0.977
45	4.362	2.867	4.949	5.349	2.700	4.712	5.191	2.874	4.949	5.351	3.019	5.188	5.494	-1.495	0.587	0.986	0.121	0.187	0.123	1.500	0.616	0.994
50	5.000	3.510	5.518	5.962	3.329	5.257	5.782	3.516	5.520	5.959	3.675	5.780	6.135	-1.490	0.518	0.962	0.135	0.207	0.142	1.496	0.558	0.972
55	5.675	4.247	6.111	6.578	4.036	5.802	6.367	4.253	6.107	6.570	4.441	6.392	6.796	-1.428	0.436	0.903	0.156	0.231	0.166	1.437	0.493	0.918
60	6.390	5.096	6.714	7.183	4.861	6.362	6.943	5.093	6.713	7.176	5.323	7.055	7.432	-1.294	0.324	0.793	0.185	0.265	0.190	1.307	0.418	0.816
65	7.146	6.070	7.323	7.779	5.781	6.952	7.481	6.077	7.318	7.770	6.334	7.708	8.074	-1.076	0.177	0.633	0.217	0.297	0.224	1.097	0.345	0.672
70	7.945	7.217	7.936	8.338	6.903	7.478	7.997	7.206	7.936	8.338	7.536	8.368	8.692	-0.728	-0.009	0.393	0.259	0.340	0.278	0.773	0.340	0.481
75	8.789	8.539	8.552	8.858	8.117	8.030	8.402	8.529	8.555	8.859	8.961	9.047	9.300	-0.250	-0.237	0.069	0.323	0.392	0.352	0.408	0.458	0.358
80	9.680	10.074	9.214	9.316	9.521	8.566	8.746	10.065	9.222	9.322	10.599	9.830	9.852	0.394	-0.466	-0.364	0.417	0.482	0.443	0.573	0.670	0.573
85	10.620	11.857	9.885	9.632	11.162	9.052	8.864	11.860	9.875	9.607	12.572	10.684	10.397	1.238	-0.735	-0.988	0.550	0.620	0.592	1.354	0.961	1.152
90	11.610	13.949	10.689	9.704	12.950	9.587	8.629	13.913	10.699	9.692	15.074	11.808	10.809	2.339	-0.921	-1.906	0.823	0.846	0.850	2.479	1.251	2.086

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.7: Results for all three approaches. Common trend intensity $c = 0.75$, treatment intensity $r = 2.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.263	0.070	0.326	1.257	-0.005	0.176	1.194	0.070	0.328	1.257	0.142	0.470	1.320	-0.193	0.062	0.994	0.057	0.117	0.051	0.202	0.132	0.995
15	0.406	0.132	0.460	1.739	0.056	0.309	1.657	0.137	0.460	1.735	0.206	0.614	1.814	-0.274	0.054	1.333	0.061	0.121	0.061	0.281	0.132	1.334
20	0.557	0.206	0.605	2.205	0.115	0.442	2.113	0.206	0.612	2.200	0.293	0.767	2.291	-0.351	0.049	1.649	0.070	0.124	0.070	0.358	0.133	1.650
25	0.716	0.298	0.757	2.651	0.193	0.566	2.546	0.298	0.766	2.654	0.403	0.931	2.757	-0.418	0.041	1.935	0.081	0.137	0.082	0.426	0.143	1.937
30	0.885	0.413	0.922	3.067	0.297	0.717	2.944	0.415	0.927	3.070	0.530	1.105	3.193	-0.472	0.037	2.182	0.092	0.150	0.096	0.481	0.154	2.184
35	1.064	0.551	1.102	3.446	0.406	0.889	3.294	0.560	1.106	3.451	0.692	1.303	3.586	-0.512	0.038	2.382	0.107	0.161	0.111	0.523	0.166	2.385
40	1.253	0.711	1.282	3.766	0.543	1.047	3.610	0.720	1.287	3.768	0.860	1.512	3.926	-0.542	0.029	2.513	0.123	0.177	0.127	0.556	0.179	2.516
45	1.454	0.903	1.470	4.013	0.720	1.230	3.826	0.914	1.471	4.012	1.071	1.728	4.189	-0.551	0.016	2.559	0.140	0.198	0.145	0.568	0.199	2.563
50	1.667	1.131	1.665	4.173	0.917	1.393	3.945	1.132	1.663	4.176	1.337	1.956	4.386	-0.536	-0.002	2.507	0.163	0.220	0.173	0.560	0.220	2.513
55	1.892	1.384	1.876	4.234	1.130	1.567	3.978	1.390	1.878	4.237	1.621	2.204	4.480	-0.508	-0.015	2.343	0.188	0.245	0.203	0.541	0.246	2.351
60	2.130	1.684	2.102	4.174	1.396	1.757	3.868	1.683	2.102	4.172	1.976	2.484	4.485	-0.446	-0.028	2.044	0.224	0.284	0.237	0.499	0.285	2.058
65	2.382	2.029	2.344	3.984	1.694	1.935	3.631	2.023	2.344	3.976	2.370	2.758	4.348	-0.353	-0.037	1.602	0.263	0.323	0.277	0.441	0.325	1.625
70	2.648	2.445	2.596	3.633	2.050	2.097	3.213	2.449	2.592	3.641	2.865	3.051	4.110	-0.204	-0.053	0.985	0.315	0.381	0.345	0.375	0.384	1.043
75	2.930	2.943	2.850	3.082	2.441	2.249	2.501	2.938	2.866	3.089	3.468	3.402	3.627	0.013	-0.080	0.152	0.395	0.447	0.436	0.395	0.454	0.462
80	3.227	3.535	3.122	2.283	2.892	2.442	1.556	3.536	3.126	2.292	4.180	3.776	2.951	0.308	-0.105	-0.944	0.515	0.556	0.548	0.600	0.565	1.091
85	3.540	4.232	3.386	1.086	3.363	2.443	0.084	4.247	3.385	1.105	5.072	4.344	2.004	0.692	-0.154	-2.454	0.690	0.721	0.733	0.977	0.736	2.561
90	3.870	5.077	3.685	-0.794	3.755	2.338	-2.067	5.105	3.694	-0.817	6.368	5.042	0.568	1.207	-0.185	-4.664	1.035	1.019	1.068	1.589	1.035	4.784

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

Table B.8: Results for all three approaches. Common trend intensity $c = 0.75$, treatment intensity $r = 7.5$

		Coeff Avg.			Coeff Q10			Coeff Q50			Coeff Q90			Bias			Standard deviation			RMSE		
τ	Target	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD	Pooled	DDiD	QDiD
10	0.790	0.202	1.094	1.784	0.126	0.922	1.721	0.202	1.091	1.784	0.276	1.262	1.847	-0.588	0.304	0.994	0.058	0.134	0.051	0.591	0.332	0.995
15	1.218	0.372	1.553	2.551	0.294	1.374	2.469	0.377	1.552	2.547	0.446	1.729	2.626	-0.846	0.335	1.333	0.062	0.135	0.061	0.848	0.361	1.334
20	1.670	0.581	2.038	3.319	0.486	1.838	3.226	0.580	2.050	3.313	0.669	2.232	3.404	-1.089	0.368	1.649	0.072	0.145	0.070	1.092	0.396	1.650
25	2.148	0.847	2.545	4.084	0.741	2.324	3.978	0.846	2.552	4.086	0.951	2.742	4.189	-1.301	0.397	1.935	0.082	0.157	0.082	1.304	0.427	1.937
30	2.655	1.181	3.070	4.837	1.057	2.836	4.714	1.183	3.079	4.840	1.303	3.281	4.964	-1.474	0.415	2.182	0.095	0.172	0.096	1.477	0.449	2.184
35	3.192	1.582	3.610	5.574	1.439	3.364	5.422	1.591	3.611	5.578	1.726	3.856	5.713	-1.609	0.419	2.382	0.108	0.183	0.111	1.613	0.457	2.385
40	3.760	2.064	4.158	6.273	1.891	3.905	6.116	2.073	4.156	6.274	2.215	4.425	6.433	-1.696	0.398	2.513	0.125	0.203	0.127	1.701	0.447	2.516
45	4.362	2.638	4.719	6.921	2.441	4.444	6.734	2.645	4.720	6.920	2.821	4.990	7.097	-1.724	0.357	2.559	0.146	0.217	0.145	1.731	0.418	2.563
50	5.000	3.320	5.287	7.507	3.106	4.972	7.278	3.322	5.281	7.510	3.530	5.601	7.719	-1.680	0.287	2.507	0.166	0.245	0.173	1.689	0.377	2.513
55	5.675	4.113	5.881	8.018	3.878	5.548	7.762	4.116	5.885	8.021	4.366	6.230	8.264	-1.562	0.205	2.343	0.196	0.267	0.204	1.574	0.337	2.351
60	6.390	5.056	6.488	8.434	4.775	6.114	8.128	5.049	6.473	8.432	5.346	6.884	8.745	-1.334	0.098	2.044	0.229	0.300	0.237	1.353	0.316	2.058
65	7.146	6.167	7.114	8.748	5.835	6.679	8.395	6.159	7.100	8.740	6.514	7.566	9.112	-0.979	-0.032	1.602	0.271	0.351	0.277	1.016	0.352	1.626
70	7.945	7.482	7.754	8.930	7.071	7.232	8.510	7.482	7.766	8.938	7.920	8.253	9.407	-0.463	-0.191	0.985	0.329	0.402	0.345	0.568	0.444	1.043
75	8.789	9.048	8.400	8.941	8.559	7.802	8.360	9.031	8.414	8.949	9.621	8.976	9.487	0.259	-0.389	0.152	0.407	0.471	0.436	0.482	0.611	0.462
80	9.680	10.922	9.093	8.736	10.283	8.347	8.010	10.920	9.086	8.745	11.612	9.827	9.404	1.242	-0.587	-0.944	0.527	0.585	0.548	1.350	0.828	1.091
85	10.620	13.151	9.800	8.166	12.211	8.897	7.164	13.193	9.793	8.185	14.014	10.712	9.083	2.531	-0.820	-2.454	0.719	0.728	0.733	2.631	1.096	2.561
90	11.610	15.845	10.595	6.946	14.450	9.254	5.673	15.882	10.615	6.923	17.190	11.888	8.308	4.235	-1.015	-4.664	1.085	1.029	1.068	4.372	1.444	4.784

Notes: Table displays the full results of the MC simulation study based on $M = 500$ MC repetitions and a sample size of 400,000. The QTT estimators for the three RIF-based approaches refer to the estimated parameters given in (3.16), (3.20), and (3.22), respectively.

B.3.2 Empirical application – full results

Table B.9: Comparison of QTT estimates from all three RIF approaches

			No controls			Full controls		
τ	$q_{\tau\|11}$	$q_{\tau\|m}$	pooled RIF	DDiD-RIF	QDiD-RIF	pooled RIF	DDiD-RIF	QDiD-RIF
5	26.606	14.374	0.556 (2.178)	2.394 (2.452)	1.975 (2.444)	−1.195 (2.278)	−0.255 (2.464)	0.157 (2.374)
10	48.956	42.921	2.880 (2.027)	2.165 (1.998)	4.863* (2.801)	−0.017 (1.938)	−0.551 (1.857)	2.050 (2.936)
15	63.309	61.229	6.504*** (2.310)	5.963*** (2.173)	2.592 (2.275)	3.229 (2.152)	3.313* (1.987)	−0.678 (2.300)
20	73.922	74.843	9.240*** (2.678)	7.273*** (2.467)	3.910* (2.077)	6.210*** (2.364)	4.694** (2.139)	1.106 (1.992)
25	83.952	86.714	14.336*** (2.261)	12.036*** (2.196)	8.659*** (2.022)	11.258*** (1.967)	9.360*** (1.875)	5.624*** (1.933)
30	94.480	98.057	18.739*** (2.031)	15.555*** (2.075)	11.288*** (1.740)	15.545*** (1.803)	12.789*** (1.753)	8.057*** (1.669)
35	104.002	108.650	19.255*** (1.720)	17.048*** (1.520)	14.360*** (1.764)	15.768*** (1.704)	14.023*** (1.506)	11.037*** (1.782)
40	112.634	118.734	18.674*** (1.564)	15.286*** (1.247)	14.157*** (2.208)	14.948*** (1.376)	12.254*** (1.185)	10.639*** (2.251)
45	120.711	128.716	15.771*** (1.462)	14.324*** (1.112)	16.913*** (2.203)	12.073*** (1.396)	11.430*** (0.996)	13.300*** (2.420)
50	128.045	139.559	15.495*** (2.120)	12.346*** (1.135)	12.923*** (2.289)	11.663*** (2.081)	9.464*** (1.064)	9.198*** (2.371)
55	136.678	151.115	16.415*** (2.383)	11.731*** (1.506)	14.665*** (2.535)	12.541*** (2.539)	8.607*** (1.430)	11.167*** (2.602)
60	145.709	163.513	11.193*** (2.502)	13.721*** (1.744)	15.837*** (2.287)	7.353*** (2.606)	10.519*** (1.821)	12.289*** (2.469)
65	156.022	177.179	9.066*** (2.899)	11.108*** (1.901)	12.528*** (2.185)	5.476* (3.031)	7.752*** (2.011)	8.612*** (2.349)
70	166.679	192.885	10.471*** (3.146)	8.183*** (2.080)	10.746*** (2.860)	6.650** (3.331)	4.816** (2.208)	6.711** (3.109)
75	179.224	209.836	5.453 (3.369)	9.372*** (3.238)	5.088 (3.175)	1.287 (3.527)	5.774* (3.329)	1.011 (3.376)
80	196.428	230.905	4.957 (4.049)	11.059*** (3.123)	3.853 (3.282)	0.564 (4.365)	6.738** (3.342)	0.048 (3.732)
85	217.515	258.218	2.781 (3.817)	7.686* (4.172)	4.580 (3.444)	−1.027 (4.394)	2.330 (4.531)	1.032 (3.769)
90	249.996	293.106	4.513 (4.426)	5.942 (6.008)	8.607* (5.185)	1.932 (4.790)	−0.055 (6.785)	5.233 (5.496)
95	307.143	355.693	5.154 (6.574)	12.187 (9.712)	11.634 (7.632)	3.693 (6.750)	7.795 (10.169)	9.294 (7.254)

Notes: Full results for coefficient estimates of terms involved in (3.28) were obtained by means of all three RIF approaches. The results under "No controls" refer to an alternative specification in which only the necessary DiD terms are included. Both the marginal quantile derived from the distribution for the treated in the post-treatment period, $q_{\tau|11}$, and the quantile derived from the pooled distribution, $q_{\tau|m}$, are given. Bootstrap standard errors (based on 1,000 repetitions, clustered at the state level) are given in parentheses. $^{***}/^{**}/^{*}$ indicate statistical significance at the a 1%/5%/10% level.
Source: Raw data provided by Tax Policy Center (2023). The data that have been used to produce this table are part of the replication data provided in Hoynes and Patel (2022).

Chapter 4

How Did De-unionization Impact the German Structure of Earnings? A Distributional Approach Using Grouped Quantile Regressions*

*This chapter is based on Rümmele, M. (2024): How Did De-unionization Impact the German Structure of Earnings? A Distributional Approach Using Grouped Quantile Regressions. *Unpublished manuscript, University of Tübingen.*

4.1 Introduction

Prior to the introduction of the federally binding minimum wage in 2015, one of the most salient labor market institutions that mediated wage setting in Germany were labor unions. In Germany, the most prominent task of unions is to represent workers in centralized bargaining with employer associations over remuneration schemes and other workplace-related matters. After successful negotiations between unions and employer associations, a collective bargaining agreement (CBA) is signed which is applicable at the sector-region level in Germany (see detailed review of the German setting below). Arguably one of the most striking labor market developments in Germany – and, in fact, in many industrialized countries – was the unprecedented decentralization of wage setting mechanisms which took place from the mid-1990s to the mid-2010s in Germany (Dustmann et al., 2014). Specifically, the coverage of workers by sectoral collective bargaining agreements decreased from 71% in 1996 to 52% in 2014 (cf., table 4.1). Given the institutional setting in Germany, the implications of this drop on the structure of earnings are not immediately obvious. On the one hand, the German setting does not imply sharp policy changes that could be utilized, e.g., in the form of union elections within firms as it is the case in the United States where unions act at the firm level (cf., DiNardo and Lee, 2004). There is furthermore not a clearly distinguishable non-union sector that could be used to infer a non-union counterfactual (Card et al., 2003, 2020).

Rather, unions in Germany act at a comparatively high level of centralization, with collective agreements mostly applying at the sector-region level. The latter implies that any firm-level consideration cannot fully account for the overall effect of union-mediated centralized wage setting mechanisms. This is due to structural peculiarities of centralized wage setting in Germany, i.e., sectoral agreements produce spillovers due to the well-known scheme of "orientation" of CBAs by employers who did not participate in collective bargaining but who apply the remuneration schemes in their firms (Bossler, 2019). Put differently, sectoral CBAs can be viewed as setting sector-specific norms and expectations about worker payments (Jäger et al., 2022). If the aim is to recover an overall effect of unions through CBA coverage in Germany, firm- or worker-level analyses thus face substantial limitations. Additionally, the majority of German data lack the necessary detail for researchers to determine if a worker is being paid according to a CBA within firms that report to follow such agreements. These institutional as well as data-driven difficulties thus require to align the methodological approach to the specific German setting.

Regarding the effect of unions/collective agreements on the earnings structure and the distribution of wages and earnings, earlier contributions established that unionization coincides with more compressed distributions and thus lower wage and earnings inequality altogether (see, e.g., Farber et al., 2021; Card et al., 2020; Firpo et al., 2018; DiNardo et al., 1996, for the US and Canada). For the German case, the stark shift in worker composition towards non-covered workers has also been found to explain much of the increase in wage inequality from the mid-1990s to the mid-2010s (e.g., Dustmann et al., 2009, 2014; Biewen and Seckler, 2019). Moreover, this remarkable compositional shift arguably bears different implications for different worker types, depending on which worker traits are the most affected by a (de-)centralized wage setting regime.

Against the backdrop of the large body of existing studies on the effect of union wage setting in Germany, this study makes the following contributions. First of all, to characterize the effect of union-mediated wage setting in Germany, we employ a novel methodological approach introduced by Chetverikov et al. (2016), which allows for examining the distributional effects of a group-level treatment using quantile regression techniques. Employing such a group-level analysis effectively models the union impact as taking effect at the sector-region level – the level at which most collective agreements apply in Germany. Thus, we fully take the German institutional setting into account. Specifically, the focus of the analysis is shifted towards the 'treatment' level rather than considering the firm or worker level. The rationale is based on the notion that both actual CBA coverage prevalence as well as spillover effects imply a relatively stronger union effect in sector-regions with a higher prevalence of sectoral collective agreements. The union effect that can be measured at the group-level is therefore conceived as being the extent to which wage setting is centralized in a given sector-region.

As a second contribution, the empirical strategy provides a rich characterization of the union effect in the German institutional setting by allowing it to differ along various dimensions. Several previous studies have considered an aggregate effect of unions or CBA coverage in Germany (e.g., Dustmann et al., 2009; Card et al., 2013; Biewen and Seckler, 2019), or considered only one specific worker trait to infer a differential effect of union wage setting (e.g., the gender wage differential as in Antonczyk et al., 2010; Bruns, 2019; Oberfichtner et al., 2020, see also literature review below). On the contrary, the approach used here allows for the estimation of differential effects along the distribution of worker productivity (the conditional distribution of earnings) as well as for a separate union effect for different worker types. The method furthermore enables the estimation of the union effect on measures of inequality both between and within worker groups. This is achieved by employing the grouped quantile regression approach introduced by Chetverikov et al. (2016) and extensions of the method proposed by Oka and Yamada (2023) and Xu et al. (2023). To the best of our knowledge, this paper is the first application of such a detailed group-level approach to examine the impact of unions on the structure of earnings.

Thirdly, we propose a novel method for modeling the unconditional effect of de-unionization by integrating the group-level approach of Chetverikov et al. (2016) with the method put forward by Machado and Mata (2005). This method allows for the simulation of outcome draws from both the factual and counterfactual unconditional earnings distributions, with the counterfactual distribution being based on a scenario where CBA coverage had not declined over time. This presents a new angle for assessing the effect of de-unionization on the unconditional distribution of earnings, an aspect not yet examined in earlier studies. Thus, we contribute new empirical evidence to the existing literature on the distributional impact of de-unionization in Germany.

To preview our main findings, our empirical analysis shows that the degree of centralization in wage setting affects earnings differentials, especially along the lines of educational attainment and gender. Using high-quality linked employer-employee data, we find that higher CBA coverage coincides with more compressed earnings differentials between workers with low and medium educational attainment in the middle and upper parts of the worker productivity distribution. Additionally, the empirical analyses suggest that higher CBA coverage coincides with a smaller male-female earnings differential almost uniformly along the distribution of worker productivity. All this translates into substantial effects of the observable drop in CBA coverage between 1996 and 2014 on between- and within-group inequality. In particular, the decline in CBA coverage accounts for much of the increased earnings differentials between low- and medium-educated worker groups during this period. While we find that the earnings differential between male and female workers has narrowed over time, the decline in CBA coverage has significantly attenuated this trend. Furthermore, we demonstrate that the reduction in CBA coverage has had a pronounced inequality-enhancing effect within the worker groups analyzed, particularly among male workers. Concerning the implied unconditional effect, our findings indicate that de-unionization has significantly contributed to the observed decline in real earnings, especially at lower to lower-middle percentiles of the unconditional earnings distribution.

The remainder of the paper is structured as follows. Section 4.2 reviews the literature. Section 4.3 provides a brief overview of how unions operate in the German institutional setting. Section 4.4 describes the data used for the empirical analysis. In sections 4.5 and 4.6, the econometric method and the empirical results are described. Finally, section 4.7 concludes.

4.2 Literature review

Unions' impact on the structure of wages and earnings

A well-documented pattern for both the US and Germany is the dramatic decline in union coverage over time (see, e.g., Stansbury and Summers (2020) for the US and Dustmann et al. (2009, 2014); Ellguth and Kohaut (2019) for Germany). To assess how labor market institutions affected the wage distribution in the US, DiNardo et al. (1996) employ a reweighting technique to infer counterfactual wage distributions that can be used to decompose changes in the wage distribution over time. Their findings suggest that much of the observable increase in wage inequality in the US during the 1980s could be accounted for by changes in the real value of the minimum wage as well as deteriorating union prevalence. More recently, Fortin et al. (2021) reconsidered the analysis by DiNardo et al. (1996) and additionally allowed for spillover effects of the institutional elements. Regarding unions, they explicitly model threat effects of union policies to the non-unionized sector, which considerably increases the explanatory power of unions in explaining changes in wage inequality in the US between 1979 and 2017. Employing Recentered Influence Function regressions (RIF, Firpo et al., 2009) as well as RIF decomposition techniques, Firpo et al. (2018) demonstrate that de-unionization had a large impact on the unconditional distribution of wages in the US, especially at the lower to middle part of the distribution over the course of a 30-year period (1980s to 2010s). Taking a long-run perspective on the impact of unions on wage inequality in the US, Collins and Niemesh (2019) and Farber et al. (2021) evaluate how unions have shaped the wage distribution over the 20th century. Collins and Niemesh (2019) establish that unions as early as in the pre-WW2 period have had a strong impact on inequality. Examining the long-run relationship from the 1930s onward, the analysis by Farber et al. (2021) confirms the role of unions as one of the major US labor market institution shaping income distributions in the 20th century – notwithstanding the fact that changes in relative skill supplies played a large role as well.

For the case of Germany, Dustmann et al. (2009) find that the tendency towards a more de-unionized economy coincided with a stark increase in lower-tail wage inequality in West Germany from the mid 1990s until 2004 (the end of their observation period), whereas the increase in upper-tail inequality was more strongly associated with changes in the composition of worker characteristics. Considering a large variety of potential factors within a RIF decomposition analysis, Biewen and Seckler (2019) find de-unionization to have been the most important factor for explaining the stark increase in wage inequality in Germany starting in the 1990s until 2010. Regarding the underlying reasons for this sharp decline in union coverage, Fitzenberger and Sommerfeld (2016) document that, during the period in which one of the largest drops in CBA coverage took place (2001-2006), this drop coincided with neither changes in characteristics nor returns to those observable worker characteristics, but is rather the result of an 'unexplained time trend.'

Unions and skill premia

Early on, Card (1996) evaluated the impact of unions on workers with different skill levels in the US, finding that low-skill workers benefited more from unions than high-skill workers. The differing selection patterns, with low- (high-)skill workers being positively (negatively) selected into unionization in the US, played an important role as well. Another implication of the analysis by Card (1996) is that differences in returns to worker characteristics are relatively more compressed for unionized workers. The extent to which the union premium varied across education and race in the US over the longer run has, among other things, been part of the contribution by Farber et al. (2021). Their findings indicate a compositional change within the group of unionized workers, which had implications for the effect of unions on, e.g., returns to education. Further, Farber et al. (2021) find that the union premium in the US over the longer run was larger for non-white workers. Beyond this, union premiums were found to be larger for the group of relatively less educated men. Regarding the importance of the unionized workforce's composition on the effect of unions on the wage structure, Card et al. (2020) find that the median US/-

Canadian union worker today is no longer male and employed in the private but female and employed in the public sector. This, in turn, changed the way unions affected labor market outcomes with the latter having led to a much larger overall effect of unions in the public than in the private sector in those countries.

Unions and gender differentials

Explicitly considering the impact of unions on the gender gap in earnings, Biasi and Sarsons (2022) show that a decentralization of wage-setting mechanisms for public school teachers in the US led to a spike in the gender pay gaps among a previously fully unionized group of workers in which no gender pay gap had been prevalent. As the timing of the ‘decentralization shock’ implies a quasi-exogenous setting, the findings by Biasi and Sarsons (2022) provide strong evidence regarding the effect of centralized wage setting mechanisms on gender-related wage differentials. Bruns (2019) highlights that the decentralization of the German labor market from the mid-1990s onwards was crucial for shaping gender inequality within firms, adding to the overall effect that women tend to work in less productive firms over the working life cycle compared to their male counterparts. Among other things, Felgueroso et al. (2008) discuss the impact of nationally or regionally collective bargaining agreements on the gender pay gap over the entire distribution of earnings, suggesting an alleviating effect of these CBAs on the gender pay gap at the bottom of the wage distribution. For the case of Germany, Antonczyk et al. (2010) provides a discussion regarding the expected direction of the union effect on the gender pay gap, arguing that the standardization of wage setting regimes, ex ante, should benefit women by alleviating gender-based discrimination. On the other hand, the incentives of unions could be centered around the median union member, which is still a male worker in most industries. In their empirical analysis they find that the de-unionization between 2001 and 2006 in Germany had the same effect for both male and female worker and thus did not affect the gender pay gap along the unconditional distribution of wages. Exploiting discrete changes in bargaining regimes at the firm-level using the same firm-level survey data as in this paper, Oberfichtner et al. (2020) find no effect of these discrete changes on the gender pay gap.

4.3 Institutional background

The German economy went through several phases during the period considered in the empirical analyses below. This also bears implications regarding the way unions act in centralized wage setting. Hence, in the following, we provide an overview of different episodes and the corresponding economic conditions during these episodes in Germany. Moreover, we briefly describe how unions operate within the specific German institutional setting of centralized wage bargaining. Lastly, we outline how the institutional setting naturally coincides with an analysis that exploits variations across sector-regions.

Economic conditions and de-unionization in Germany

Given the varying economic conditions in Germany from the mid-1990s to 2014, it is important to consider the episode-specific effects of shifts in CBA coverage on the earnings structure in our empirical analysis. Allowing for this temporal heterogeneity is potentially insightful as unions effectively need to take the overall economic conditions into account when bargaining over new collective agreements and need to react accordingly, e.g., by means of wage restraints to mitigate employment effects if employment prospects are weak in economic downturns or, vice versa, to bargain more aggressively in times of a roaring economy when larger rents are to be shared (see, e.g., Dribbusch et al., 2017, for a review on various phases of union impact in Germany). Furthermore, episodic events and the various elements of centralized wage setting (i.e., union density, CBA applicability, membership in employer associations, opening clauses) are intertwined to a certain extent as, e.g., argued in Jäger et al. (2022).

The first episode we consider comprises the years 1996 to 2001 and essentially represents the years following German reunification. These years were characterized by a sharp restructuring of the labor market and an increasing internationalization of the economy following the fall of the Iron Curtain. During the 1990s, East German firms were pressured to align wages with the prevailing rates in the Western part of the country, where wages were predominantly stipulated by collective bargaining (Dustmann et al., 2014). These wage rates often proved to be too high for East German firms which prompted them to defect from the system of centralized wage setting. Moreover, the de-unionization that took place in the new federal states spilled over to the Western part of the country, leading to a decline of unions' impact in the German labor market altogether (Card et al., 2013). Moreover, the growing internationalization of the economy in the post re-unification period has been found to have triggered some off-shoring tendencies that weakened the stance of unions in wage negotiations as well (Dauth et al., 2014; Jäger et al., 2022).

The second episode, covering 2002 to 2007, describes a period when Germany was referred to as the 'sick man of Europe' (Dustmann et al., 2014). During this time, economic and labor market conditions were bleak, ultimately leading to one of the largest labor market reforms in the history of the Federal Republic, the *Hartz IV* reforms, in 2005. Overall, this period can be characterized as one of labor market insecurity with weak to moderate economic growth.

Finally, the third period (2008 to 2014) is characterized by the recovery from the global financial crisis with subsequent strong economic growth and, in turn, prosperous labor market conditions. Even though being hit by a strong economic downturn with large drops in (real) GDP in 2009, the subsequent economic development were found to be prosperous, with record-low unemployment rates in the years after 2010 (Blömer et al., 2015).

How do unions impact the wage setting in Germany?

The term wage or earnings structure, as used in this paper, is understood as the relationship between observable worker characteristics and earnings, i.e., the extent to which earnings differ systematically between different worker types (similar to, e.g., Card et al., 2003). Beyond that, the structure of earnings within a sector-region is assumed to be shaped by certain institutional elements (CBAs in the case at hand). This adds another layer to be taken into account when unraveling how observable worker traits are remunerated. In Germany, the impact of unions on the wage-setting process manifests itself most prominently through the outcome of centralized bargaining between unions and employers' association with the latter comprising employers that act within a given industry-region.[1] The additional regional disaggregation of sectoral bargaining can be viewed against the background of accommodating regional differences that are present within a given sector (Fitzenberger and Sommerfeld, 2016). In particular, this regional component prevents

[1] Jäger et al. (2021) provide an overview of the various aspects of co-determination in the German labor market beyond CBAs.

sector-specific agreements to be too rigid. Specifically, a collective agreement that dictates the wage setting for a sector across all regions might stipulate wage rates that lead to higher unemployment in a relatively less productive region (Jäger et al., 2022).

The agreed terms, de jure, only apply to union members. This means that employers are not legally obligated to pay the same wage to comparable non-union members. Nevertheless, employers often apply the same negotiated terms to both union and non-union members, which frequently results in the de facto applicability of the CBA to all of an establishment's employees. This also relates to the fact that the German constitution stipulates that neither of the two parties must be forced to enter an employer association or a union, respectively (*negative Koalitionsfreiheit*). The latter, in principle, formulates some limit on what can be granted in premiums that would be exclusive to union members.

This expansion of the CBA terms to non-union members could be considered as a strategy on the part of employers to prevent workers joining the union, essentially eradicating any immediate financial benefit of being a union member.[2] In turn, this often implies a large coverage rate within a firm or sector, while the union density is comparatively small. As pointed out in Fitzenberger et al. (2013), union density (or membership rate) can be considered as a measure of union power which impacts both the coverage rate and whether the terms in the collective agreements are more or less in favor of the union members. On the other hand, the coverage rate can be seen as a measure of the actual applicability of the resulting CBAs in the respective sub-labor markets. It thus provides an indicator of the extent of wage-setting centralization per sector-region.

Why employing a group-level analysis?

In the empirical analysis below, we consider the union impact to be a *group-level* impact. This is due to the specific institutional setting in Germany and the available data. Ideally, data covering many periods would enable researchers to exploit quasi-experimental discrete changes in workers' coverage status to identify the effect of CBA coverage on wages or other derived parameters. However, such data are not available for Germany

[2] However, as pointed out by Hirsch et al. (2022), this employer rationale seem to be in retreat in recent years with less and less non-union members being paid according to the CBA in covered firms.

due to both data availability and institutional reasons.[3] The database that comes closest to this ideal situation are linked employer-employee data – the *LIAB* (Ruf et al., 2021) – provided by the German Institute for Employment Research (*IAB*). The *LIAB* dataset is used for the empirical analyses below, as it contains detailed administrative worker-level data and extensive firm-level survey data (see section 4.4 for a detailed description). However, a remaining issue is that information on CBA coverage is only provided at the firm level, making it impossible to exploit discrete changes in worker-level coverage status.

In principle, the *LIAB* would also allow for some sort of firm-level analysis in which one could, for example, exploit the effect of changes in firm-level coverage status (e.g., Oberfichtner et al., 2020). However, such a firm-level analysis suffers from several shortcomings as pointed out by Jäger et al. (2022). Beside the obvious concern that such changes are not quasi-random but potentially coincide with episodes of economic hardships of the respective firms, aspects such as panel attrition and few worker observations for smaller firms complicate the empirical analysis even more. In this paper's approach, these issues are circumvented by 'zooming out' to the sector-region level.

However, we argue that the primary reason for choosing a group-level analysis is due to the German institutional setting. In particular, the impact of the CBA coverage prevalence on the structure of earnings and wages manifests itself not only through a collective agreement that is reported to be applied in a given firm. That is to say that CBAs actually resonate beyond the firms that 'actively' apply them due to the phenomenon of 'orientation' (Bossler, 2019) as many employers who did not actively participate in collective bargaining apply the remuneration schemes of the relevant sectoral collective agreement in their companies as well. Similarly, Oberfichtner and Schnabel (2019) show that the voluntary orientation of uncovered firms played an increasing role during the last two decades, somewhat alleviating the overall effect of de-unionization. This, in turn, highlights the fact that a firm reporting not to actively adhere to a CBA does not rule out the possibility that they still voluntarily follow schemes dictated by CBAs.

CBAs can thus be viewed as setting norms in worker remunerations outside their actual

[3]The only data source in Germany that contains detailed information on the *individual* application of collective agreements is the German Structure of Earnings Survey (*GSES*). However, a structural break in the information on individual coverage occurred in 2014, making comparisons over time, including 2014, problematic. Additionally, the data have only been available every four years as a repeated cross-section, preventing effective analysis of variations within workers or firms over time.

scope (Jäger et al., 2022). Analyzing the effect of CBAs by exploiting group-level variation reflects the fact that shifts in the prevalence of this particular labor market institution have more far-reaching consequences than those detectable through firm-level analyses. To carry out the group-level analyses, the group-specific share of employees working in covered firms can be considered a good indicative measure of the scope of 'unionization' in Germany. In other words, the effect measured in the empirical analysis is the effect that corresponds to shifts in the degree of wage-setting centralization resulting from shifts in the prevalence of CBA coverage. A drop in CBA coverage in a sector-region is perceived to weaken the 'norm-setting' power of union wage setting, compared to sector-regions where CBA coverage remains higher.

As to the underlying mechanisms behind CBA-induced earnings adjustments, the prevalence of CBA coverage can be viewed as a mediating force in the labor market (Autor and Katz, 1999). This force can be understood as translating market-driven forces (e.g., due to technology-augmented shifts in the supply/demand for particular types of workers, or shifts in product markets due to changes in consumer tastes or technological changes) into actual earnings adjustments (cf., Bound and Johnson, 1992; Autor and Katz, 1999, for a more formal representation).

4.4 Data and descriptives

4.4.1 The LIAB

Our econometric approach involves the consideration of sector-region-specific returns to worker observables over time. Hence, the data need to fulfill several requirements. Firstly, data must be available over a long period of time to examine the drop that took place from the mid-1990s to the mid-2010s. Secondly, one needs enough information on the group level to actually estimate these parameters, i.e., there needs to be a sufficient amount of worker-level observations *within each sub-labor market*. Lastly, the data must contain detailed information on both earnings and other work-related variables, including CBA coverage.

For the empirical analysis, the *LIAB Cross Sectional Model* provided by the *IAB* is used

(*LIAB QM2 9319*, see Ruf et al., 2021, for further information). Concretely, the *LIAB* is a linked employer-employee data set that links administrative worker-level information from the *Integrated Employment Biographies* (*IEB*) with data from a yearly establishment survey (*IAB Establishment Panel*, *BP*). The firms in the *BP* are sampled from the universe of German firms that employed at least one employee subject to social security contributions one year prior to the survey's cutoff day (June 30th).[4] Data on employed individuals stem from employers' social security notifications (Employee History, *BeH*). The administrative data on workers do not comprise civil servants and self-employed individuals.

Crucially, the *LIAB* is available for each and every year from 1996 to 2014 in a consistent manner. While the *LIAB* in principle provides data from 1993 to 2019, only the years 1996 to 2014 are used because information on CBA coverage is consistently available only from 1996 onward. To focus on the effect on earnings structures due to shifts in CBA coverage, the last year included in the analyses is 2014. This is because 2014 is the final year unaffected by the federal minimum wage introduced in 2015, which significantly impacted the wage and earnings distribution in Germany (Bossler and Schank, 2023; Biewen et al., 2022). The sampling procedure implies that the *LIAB* contains more worker-level observations from large firms, federal states, or industries. Hence, to ensure that the results are representative, cross-sectional weights provided in the *LIAB* are used throughout the analyses.

4.4.2 Variables in the analysis

Worker-level variables

Worker observations in the *LIAB* are organized in social security notification spells. Thus, data on workers' earnings refer to the remuneration a given employee received within a given spell. Since no working hours information is available from social security notifications beside the information on whether an individual worked full- or part-time, the earnings information that will be used below is derived from *daily earnings*, i.e., total spell remuneration divided by the spell's number of days. The dependent variable for modeling the earnings structure below is the log of real daily earnings using the CPI to deflate wages.

[4]For more information on the *BP* see Bechmann et al. (2021).

As the data are derived from administrative social security notifications, earnings are top-coded in the *LIAB* due to the social security contribution ceiling. The econometric approach below, however, allows to flexibly take the issue of right-censoring into account when carrying out the worker-level analyses.

The structure of earnings is modeled in terms of workers' education, experience, and gender. Several reasons justify the selection of these three dimensions. First, as suggested by Oka and Yamada (2023), education, age, and gender are characteristics that are predetermined before an individual enters the labor market and are commonly used in modeling wage and earnings structures. Moreover, as discussed above, whether a more centralized wage-setting regime affects the gender pay gap is of obvious interest. Furthermore, when evaluating the effect of unions and CBAs on differently-skilled workers, educational attainment and age are often used to differentiate skill levels (Card et al., 2003, 2020).

Regarding actual labor market experience, it is important to note that while this information is provided in the *LIAB* as a generated biographical element, it is left-censored at 1990 for individuals from East Germany and at 1975 for individuals from West Germany. This, in turn, could lead to a distorted picture of labor market experience, especially in the early years of the considered period. Hence, age is used as a proxy for potential labor market experience. To account for potential non-linearities in the return to experience, four age categories (aged 18-25, 26-35, 36-50, and 51-64) are used for modeling the earnings structure.

The information on educational attainment in the *LIAB* used for the analyses combines information on formal education and vocational training. Since comparatively small groups will be considered for modeling the group-specific earnings structures, the original categories of educational attainment provided in the *LIAB* are aggregated. Ultimately, three educational categories are used in the empirical analysis, i.e., no vocational training/other, vocational training, and university/university of applied sciences graduates.[5]

[5] For a detailed description of the original categories of educational attainment, see Ruf et al. (2021). Thomsen et al. (2018) further describe data quality problems due to structural breaks in the *BeH* in 2011. The educational variable used in this work is developed by Fitzenberger et al. (2006) and is further described in Thomsen et al. (2018).

Group-level variables

As described above, unions are assumed to influence the structure of earnings within the German context through the prevalence of CBAs in specific sector-regions. Specifically, the impact of a more or less centralized wage-setting regime in a given sector-region is conceptualized in terms of the fraction of individuals employed in firms that were reported to apply a CBA. As the empirical analysis focuses on the group in the form of the sector-region, only the *sectoral* and not the firm-level collective agreements (so-called *Haustarifverträge*) are considered when calculating the group-specific prevalence of CBAs.

The first element of the group definition is the federal state in which a given employer operates. This information is available in the *BP*. Since some federal states contain only comparatively few observations, smaller federal states were lumped together (see table 4.1 for the resulting regional classification). Regarding the sectoral element of the group, the sector classification that is used for stratifying the sampling of firms is used (Bechmann et al., 2021), resulting in 19 distinct sectors (see table 4.1).

For the final sample, several sample selection restrictions were imposed. The final sample for the analyses consists of full-time workers being employed subject to social security contributions which were at least 18 and younger than 65 years old. We choose to consider only full-time employees because social security notifications lack precise information on working hours beyond the full-time/part-time indicator. It is thus difficult to compare full-time and part-time individuals, especially when it comes to analyses that involve earnings and inequality due to the substantial dispersion of working hours among part-time workers (see also the discussion in Dustmann et al., 2009; Card et al., 2013). Furthermore, observations with zero daily earnings are dropped. If there are multiple observations for a given worker, the observation with the highest daily earnings is kept. In addition, certain worker groups were excluded from the analysis. Concretely, members of the armed forces as well as individuals being affiliated to the sectors “extraterritorial organizations” or “activities of households as employers” were excluded.

After all sample selection steps, the raw worker-level observations per year range from a minimum of 1,094,048 in 2014 to a maximum of 2,014,225 in 1996. The number of firms range from a minimum of 7,809 in 1996 to a maximum of 14,256 in 2001.

4.4.3 Descriptives

Table 4.1 contains the fractions of workers along the discussed worker- and group-level dimensions. Moreover, the prevalence of CBA coverage is given both for all sectors considered as well as for the economy altogether. Since a total of 19 years are considered, we provide the fractions for four selected years to maintain readability while still allowing for an examination of how the compositions have evolved over time.

Table 4.1: Descriptive statistics: Worker-level and group-level characteristics

	1996	2002	2008	2014
Worker characteristics (shares)				
Sex				
Male	0.651	0.649	0.667	0.691
Education				
No voc. training / other	0.104	0.096	0.083	0.077
Vocational training	0.787	0.780	0.768	0.739
University / University of applied sciences	0.109	0.124	0.148	0.184
Age groups				
18-25	0.095	0.089	0.082	0.077
26-35	0.322	0.251	0.211	0.227
36-50	0.390	0.460	0.468	0.382
51-64	0.193	0.200	0.240	0.314
Group characteristics				
Federal states (shares)				
Schleswig-Holstein/Hamburg	0.051	0.058	0.058	0.059
Lower Saxony/Bremen	0.103	0.094	0.095	0.098
North Rhine-Westphalia	0.205	0.210	0.212	0.217
Hesse	0.074	0.080	0.079	0.080
Rhineland-Palatinate/Saarland	0.055	0.057	0.055	0.054
Baden-Wuerttemberg	0.124	0.139	0.141	0.138

Continued on next page

Table 4.1 – continued from previous page

	1996	2002	2008	2014
Bavaria	0.143	0.155	0.171	0.169
Berlin	0.045	0.051	0.040	0.040
Mecklenburg-Western Pomerania/Brandenburg	0.060	0.047	0.043	0.043
Saxony	0.061	0.051	0.051	0.049
Saxony-Anhalt/Thuringia	0.077	0.058	0.054	0.053
Sectors (shares)				
Agriculture/forestry	0.007	0.009	0.009	0.009
Mining, energy, waste disposal, water supply	0.030	0.023	0.018	0.023
(Luxury) foods	0.036	0.027	0.026	0.024
Manufacturing of consumer goods	0.030	0.027	0.024	0.017
Manufacturing of production goods	0.055	0.056	0.058	0.060
Manufacturing of capital goods	0.179	0.173	0.188	0.173
Construction	0.101	0.075	0.065	0.071
Wholsale, automobile trade/repair	0.081	0.090	0.084	0.081
Retail	0.060	0.058	0.055	0.049
Transport, warehousing	0.045	0.056	0.065	0.062
Information, communication	0.022	0.021	0.024	0.032
Hospitality sector	0.024	0.026	0.027	0.024
Finance/insurance services	0.037	0.040	0.033	0.029
Economic/scientific/self-employed services	0.070	0.109	0.133	0.157
Education, teaching	0.035	0.027	0.024	0.025
Healthcare, social services	0.081	0.081	0.085	0.080
Other services	0.015	0.017	0.015	0.017
Non-profit organizations	0.017	0.013	0.015	0.013
Public administration	0.074	0.073	0.052	0.053
CBA coverage across sectors (mean)				
Agriculture/forestry	0.323	0.361	0.291	0.311
Mining, energy, waste disposal, water supply	0.775	0.603	0.666	0.658
(Luxury) foods	0.772	0.728	0.528	0.511
Manufacturing of consumer goods	0.716	0.688	0.483	0.429
Manufacturing of production goods	0.780	0.719	0.594	0.615
Manufacturing of capital goods	0.781	0.617	0.554	0.516

Continued on next page

Table 4.1 – continued from previous page

	1996	2002	2008	2014
Construction	0.748	0.694	0.653	0.681
Wholsale, automobile trade/repair	0.630	0.620	0.408	0.403
Retail	0.717	0.635	0.462	0.331
Transport, warehousing	0.642	0.555	0.402	0.377
Information, communication	0.227	0.239	0.216	0.145
Hospitality sector	0.596	0.539	0.631	0.483
Finance/insurance services	0.903	0.878	0.823	0.765
Economic/scientific/self-employed services	0.409	0.312	0.477	0.504
Education, teaching	0.813	0.616	0.576	0.568
Healthcare, social services	0.716	0.614	0.517	0.490
Other services	0.628	0.530	0.461	0.452
Non-profit organizations	0.426	0.461	0.438	0.514
Public administration	0.912	0.707	0.784	0.876
CBA coverage (mean, overall economy)	0.709	0.599	0.532	0.517
N	2,017,225	1,771,475	1,380,027	1,094,048

Notes: Table shows shares of all categorical variables as well as the economy-wide and sector-specific CBA coverage share for the four years 1996, 2002, 2008, and 2014. Federal states were lumped together to ensure large enough groups for the within-group analyses. The sample size refers to the nominal sample size of the working sample that is used for the empirical analyses. *LIAB* sampling weights are applied in all calculations.
Data: LIAB QM2 9319, own calculations.

Regarding the evolution of workers in terms of gender, educational attainment, and experience, the figures in table 4.1 show that the full-time workforce has become less female, more educated, and older. Concretely, the share of male workers increased from roughly 65% in 1996 to 69% in 2014. As to educational attainments, while 11% of workers in 1996 had an university education, this figures increased to 18% in 2014. Meanwhile, the shares of workers with vocational (no formal) training decreased from 79% (10%) in 1996 to 74% (8%) in 2014, respectively.

Concerning the group-level characteristics, the fraction of workers from East German regions declined, whereas the share of workers in West German federal states either increased or remained stable. For example, while roughly 8% of the German workforce in 1996 were employed in the federal states of Saxony-Anhalt and Thuringia, this share droped to 5% in 2014. On the contrary, the share of Bavarian workers increased by 2.6%-points from 14.3% in 1996 to 16.9% in 2014. The sectoral composition, on the other hand, remained fairly stable over time.

As to the temporal evolution of CBA coverage rates across sectors, figure 4.1 reveals substantial temporal and cross-sectoral heterogeneities. Considering the temporal heterogeneity, figure 4.1 displays a remarkable drop in CBA coverage over time across almost all sectors. Overall, while there is some heterogeneity in the extent of the decline, the overall tendency is clearly downward trending for the majority of the sectors, with only two of the 19 sectors experiencing minor increases in CBA coverage prevalence. The overall economy saw a drop in sectoral CBA coverage by 19.2%-points between 1996 and 2014, whereas some sectors experienced an even starker drop, exemplified by the retail sector, in which the CBA coverage dropped by 39%-points.

Figure 4.1: Change of CBA coverage relative to 1996 across German sectors/overall economy

Notes: The heatplot shows the percentage-points change of sectoral CBA coverage both for the overall economy (first row) as well as across German sectors (second to last row) relative to 1996, the base year. Red (blue) coloring indicates higher (lower) collective bargaining coverage compared to the base year. The darker the color, the larger the change. The CBA shares were computed using the *LIAB* sampling weights and thus are representative for the overall economy and sectors, respectively. Absolute values of CBA coverage shares for the years 1996, 2002, 2008, and 2014 can be found in table 4.1.
Source: LIAB QM2 9319, own calculations.

4.5 Methodology

In the following, the econometric methodology is outlined. In general, we closely follow the approach introduced by Oka and Yamada (2023) who evaluate the impact of changing levels of real minimum wages on the US wage structure.

4.5.1 Two-stage group level analysis

The aim of the econometric analysis is to estimate the effect of shifts in the prevalence of CBA coverage on the structure of earnings. For this purpose, the two-stage group-level quantile regression approach proposed by Chetverikov et al. (2016) is used. In brief, this methodology involves first modeling group-specific earnings structures for all sector-regions and time periods using conditional quantile regression models. The second step subsequently evaluates how changes in group-level CBA coverage affect various components of the earnings structures modeled in the first step.

Specifically, we observe a repeated cross section of workers, denoted by $i = 1, \dots, N_{jt}$, with N_{jt} being the total number of individual worker observations in sector-region j at point in time t. In the first stage of the two-step procedure by Chetverikov et al. (2016), we model the τth conditional quantile given worker-level covariates. Thus, the effects are allowed to vary for observationally equivalent workers, which, following Koenker and Bassett Jr. (1978), is formalized in the following linear conditional quantile regression model:

$$Q_{jt}(\tau \,|\, z_{ijt}) = z_{ijt}^{\intercal} \gamma_{jt}(\tau), \quad \tau \in (0,1), \tag{4.1}$$

with $Q_{jt}(\tau \mid z_{ijt})$ being the τth conditional quantile of log real earnings for workers with observed characteristics z_{ijt} in sector-region j in year t, with $t \in \mathscr{Y} = \{1996, 1997, \dots, 2014\}$. z_{ijt} is the K-dimensional vector of observable worker characteristics comprising educational attainment, potential labor market experience, gender, and a constant term. For later reference, the total number of sector-regions is referred to as J and the total number of years as T.

In the second step of the procedure outlined by Chetverikov et al. (2016), we model heterogeneities in earnings structures using group-level variables with a linear panel regression model. Most importantly, the second-stage analysis serves to disentangle the extent to which shifts in earnings structures, as modeled in (4.1), can be explained by ceteris paribus shifts in CBA coverage. In practice, we compile a panel of estimated first-stage returns from quantile regressions conducted across all sub-labor markets J and all periods T. Ultimately, the second-stage panel data regression for the kth component of the earnings structure reads

$$\gamma_{k,jt}(\tau) = u_{jt}\beta_k(\tau) + x_{jt}^{\intercal}\theta_k(\tau) + \varepsilon_{k,jt}(\tau), \quad k = 1,\ldots,K,\ \tau \in (0,1), \tag{4.2}$$

where the Q-dimensional vector x_{jt} contains a constant, a full set of sector, region, and time dummies, as well as – following Oka and Yamada (2023) – a linear trend for the involved sector and region categories to control for a potentially confounding trend in the return to the kth worker characteristic over time. u_{jt} is the CBA coverage prevalence as described above such that $\beta_k(\tau)$ gives the ceteris paribus effect of shifts in the extent of wage centralization on individual aspects of the earnings structure. Lastly, $\varepsilon_{k,jt}(\tau)$ is an idiosyncratic group-specific error term satisfying $\mathbb{E}[\varepsilon_{k,jt}(\tau)\,|\,u_{jt}, x_{jt}] = 0$ for all τ.

Having specified a linear model that describes how elements of the earnings structure vary with shifts in CBA coverage, it is crucial to discuss the interpretability of these effects. Specifically, the isolated effects of CBA coverage on the elements of the earnings structure, as detailed in (4.2), do not correspond to causal effects in a general equilibrium sense. Intuitively, even though within-group spillover effects from the covered to the uncovered sector are acknowledged, these effects can at most be interpreted as partial equilibrium effects. This is because the specification in (4.2) corresponds to considering a slight shift in CBA coverage while holding the sectoral and regional composition fixed.

Against the backdrop of the between- and within-group inequality analyses below, it is furthermore useful to describe the second-stage panel data model from (4.2) in terms of a system of K equations:

$$\underset{(K\times 1)}{\gamma_{jt}(\tau)} = \underset{(K\times K)}{\left(u_{jt}\mathbf{I}_K\right)}\underset{(K\times 1)}{\beta(\tau)} + \underset{(K\times KQ)}{\left(\mathbf{I}_K \otimes x_{jt}^{\intercal}\right)}\underset{(KQ\times 1)}{\boldsymbol{\theta}(\tau)} + \underset{(K\times 1)}{\varepsilon_{jt}(\tau)}, \quad \tau \in (0,1). \tag{4.3}$$

Conceptually, this allows for a joint consideration of the entire earnings structure as described by all its K elements.

Lastly, to allow for heterogeneity across several episodes as well, a slightly augmented version of (4.2) can be considered:

$$\gamma_{k,jt}(\tau) = u_{jt}\beta_{0k}(\tau) + (D_{1t} \times u_{jt})\beta_{1k}(\tau) + (D_{2t} \times u_{jt})\beta_{2k}(\tau) + x_{jt}^{\intercal}\theta_k(\tau) + \varepsilon_{k,jt}(\tau), \quad (4.4)$$

with $D_{1t} \equiv \mathbb{1}[2002 \leq year_t \leq 2007]$, and $D_{2t} \equiv \mathbb{1}[2008 \leq year_t \leq 2014]$ being the respective episode indicators.

This novel approach for characterizing the impact of collective wage setting in Germany differs from existing decomposition approaches, which are based on the assumption that shifts in the workforce composition do not affect the wage structure and vice versa. In particular, these decomposition approaches treat the phenomenon of de-unionization as a composition effect that is evaluated under a given wage structure.[6] On the contrary, the group-level analysis employed in this study explicitly models the effect of the compositional shift towards a more de-unionized economy *on various elements of the structure of earnings*, thereby explicitly considering the interaction between the compositional changes and the structure of earnings. This, in turn, allows for an examination of the impact of more decentralized wage setting regimes through the lens of various aspects of the earnings structure.

Estimation

As the worker-level data in the *LIAB* originate from administrative social security records, the information on earnings is censored from above by the social security contribution ceiling. This needs to be taken into account when estimating the first-stage parameters from (4.1).[7] Approaches aiming at overcoming censoring issues in the context of quantile regression techniques can intuitively be thought of as focusing only on those observations for which "censoring does not matter". This idea has been formulated in some early

[6]Frequently employed decomposition techniques that allow for a decomposition of overall shifts into a composition and a wage structure effect are the hybrid RIF-decomposition technique (Firpo et al., 2018) and the semi-parametric reweighting approach by DiNardo et al. (1996).

[7]Note that, even though the overall fraction of individuals being affected by this ceiling is comparably small, the fraction can be quite large when considering some particular worker cells (Dauth and Eppelsheimer, 2020).

contributions such as Powell (1984, 1986), and Buchinsky (1994). Crucially, a censored quantile regression approach avoids the need to impute data for censored observations.

Technically, top-coding implies that one faces a (true) latent earnings outcome, w^*_{ijt}, for which only $w_{ijt} = \min(w^*_{ijt}, c_{jt})$ is observable, i.e., the observed value is either the censoring threshold or the actual value.[8] Following Powell (1986), the censored analogy to the model specified in (4.1) arises as follows:

$$Q_{jt}(\tau \,|\, z_{ijt}) = \min\left(c_{jt}, z_{ijt}^{\intercal}\gamma^c_{jt}(\tau)\right),$$

with the c superscript indicating that parameters relate to the non-censored values of the dependent variable. Conceptually, as proposed by Powell (1986), the estimation of involved parameters only takes into account those covariate cells for which the censoring point does not play a role, i.e.,

$$\widehat{\gamma}^c_{jt}(\tau) = \arg\min_{d\in\mathbb{R}^K} \sum_{i=1}^{N_{jt}} \mathbb{1}[z_{ijt}^{\intercal}\widehat{\gamma}^c_{jt} < c_{jt}]\, \rho_\tau\left(w_{ijt} - z_{ijt}^{\intercal}d\right), \tag{4.5}$$

with $\rho_\tau(e) = [(\tau - \mathbb{1}[e \leq 0]) \cdot e]$ being the check function. Intuitively, the problem of censoring in a quantile regression framework can be addressed by identifying the covariate cells for which censoring does not apply and using only those observations for estimating the parameters. Technically, the aim is to overcome the implied non-linearity in (4.5) by focusing on cases where the indicator function equals one (Angrist and Pischke, 2008). Notably, the optimization in (4.5) is no longer a simple linear programming problem but requires further iterative steps to actually identify 'valid' observations.[9]

To estimate the first-stage group-specific quantile regression coefficients in (4.1), the three-step algorithm approach proposed by Chernozhukov and Hong (2002) is employed. For the case at hand, this culminates in the following estimator for the group-specific

[8]Since the social security contribution ceiling changes from year to year and further differs between East and West Germany, c_{jt} holds both a j and a t index. See for example Gartner (2005) or Dauth and Eppelsheimer (2020) for further information on the censoring issue present in social security data.

[9]Earlier approaches involve iterative linear programming algorithms to overcome censoring in the context of quantile regressions are proposed by Buchinsky (1994) and Fitzenberger (1997).

coefficients:

$$\widehat{\gamma}_{jt}^{1}(\tau) = \arg\min_{d\in\mathbb{R}^K} \sum_{i\in J_{1,jt}} \rho_\tau\left(w_{ijt} - z_{ijt}^{\intercal} d\right), \tag{4.6}$$

where $J_{1,jt}$ refers to the final set of observations resulting from the three-step procedure of Chernozhukov and Hong (2002). The three-step algorithm is described in detail in appendix C.1. For estimating the parameters, worker-level weights provided in the *LIAB* are used. Being a linear group-level panel data model, the group-level coefficients of the second stage (equation (4.2)) are estimated by weighted least squares using the aggregated *LIAB* weights. These aggregated weights are described in more detail in appendix C.2.

As to the computation of standard errors for the coefficients obtained in the second stage group-level analysis, we rely on the asymptotic results developed in Chetverikov et al. (2016). They derived a condition under which the first-stage sampling error becomes asymptotically negligible. These conditions relate to a sufficient number of within-group (worker) observations relative to the number of groups in the second stage. When implementing the first stage censored quantile regressions, it is thus ensured that each group contains at least 1,000 worker observations.[10] Crucially, Chetverikov et al. (2016) also establish validity for clustered applications, i.e., when estimating standard errors for the group-level parameters in (4.2) and (4.4), clustering at the level of sector-regions is taken into account.

4.5.2 Impact of CBA coverage drop on earnings inequality

Oka and Yamada (2023) and Xu et al. (2023) extended the two-stage approach of Chetverikov et al. (2016) so that the effect of observable shifts in a group-level variable on certain between- and within-inequality measures can be investigated. The subsequent subsections introduce the conceptual framework, while appendix C.3 provides additional remarks regarding the estimation of the introduced parameters.

[10]As to the group dimension, there are, in total, 3,971 group (11 regions × 19 sectors × 19 years).

Factual and implied counterfactual quantiles

In what follows, we apply the second-stage procedure described above to model the aggregate impact of de-unionization between 1996 and 2014 at different relative positions of the conditional earnings distribution to construct counterfactual earnings structures. First, note that the second-stage group-level analysis corresponds to a linear model for the conditional expected value of the first-stage returns, given group-level covariates. Stacking all K first-stage worker returns over each other as in (4.3), this can be expressed as:

$$\mathbb{E}[\underset{(K\times1)}{\gamma_{jt}(\tau)} \mid u_{jt}, x_{jt}] = \underset{(K\times K)}{(u_{jt}\mathbf{I}_K)}\underset{(K\times1)}{\beta(\tau)} + \left(\underset{(K\times K)}{\mathbf{I}_K} \otimes \underset{(1\times Q)}{x_{jt}^{\intercal}}\right)\underset{(KQ\times1)}{\boldsymbol{\theta}(\tau)}\ . \tag{4.7}$$

The formulation in (4.7) can be used to derive a counterfactually-implied vector of returns to worker observables. The counterfactual incorporates the thought experiment of holding the CBA coverage constant at the level observed in 1996, denoted as $\bar{u}_{jt_0}$, and examining the implied returns in 2014 (t_1) if the CBA coverage had remained at this counterfactual level rather than its actual level. Formally:

$$\mathbb{E}[\underset{(K\times1)}{\gamma_{jt_1}^{cf}(\tau)} \mid \bar{u}_{jt_0}, x_{jt_1}] = \underset{(K\times K)}{(\bar{u}_{jt_0}\mathbf{I}_K)}\underset{(K\times1)}{\beta(\tau)} + \left(\underset{(K\times K)}{\mathbf{I}_K} \otimes \underset{(1\times Q)}{x_{jt_1}^{\intercal}}\right)\underset{(KQ\times1)}{\boldsymbol{\theta}(\tau)}\ . \tag{4.8}$$

Hence, considering the difference of (4.7) and (4.8), the isolated effect of the CBA coverage drop in t_1 can be expressed as

$$\mathbb{E}[\underset{(K\times1)}{\gamma_{jt_1}(\tau)} \mid u_{jt_1}, x_{jt_1}] - \mathbb{E}[\underset{(K\times1)}{\gamma_{jt_1}^{cf}(\tau)} \mid \bar{u}_{jt_0}, x_{jt_1}] = \left[\left(u_{jt_1} - \bar{u}_{jt_0}\right)\mathbf{I}_K\right]\beta(\tau). \tag{4.9}$$

Furthermore, by the law of total expectations, an aggregated version of (4.9) can be formulated:

$$\mathbb{E}[\mathbb{E}[\underset{(K\times1)}{\gamma_{jt_1}(\tau)} \mid u_{jt_1}, x_{jt_1}] - \mathbb{E}[\underset{(K\times1)}{\gamma_{jt_1}^{cf}(\tau)} \mid \bar{u}_{jt_0}, x_{jt_1}]] = \left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}]\right)\mathbf{I}_K\right]\beta(\tau). \tag{4.10}$$

The formulation in (4.10) can be used to examine the isolated effect of the CBA coverage drop on a specific worker-type defined by a given K-dimensional vector $\bar{z}$. This is equivalent to the expected values (Oka and Yamada, 2023, call this 'national means') of the difference between the factual and the counterfactual type-$\bar{z}$ conditional quantiles, i.e.,

$$\Delta^{f,cf}(\tau\,|\,\bar{z}) \equiv \bar{z}^{\intercal}\left(\left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}]\right)\mathbf{I}_K\right]\beta(\tau)\right) = \mathbb{E}\left[Q_{jt_1}(\tau\,|\,\bar{z}) - Q^{cf}_{jt_1}(\tau\,|\,\bar{z})\right]. \quad (4.11)$$

To understand how the second equality in (4.11) arises, reconsider the functional form introduced in (4.1), and note that the implied counterfactual quantile reads as $Q^{cf}_{jt}(\tau\,|\,\bar{z}) = \bar{z}^{\intercal}\gamma^{cf}_{jt_1}$, i.e.,

$$\mathbb{E}\left[Q_{jt_1}(\tau\,|\,\bar{z}) - Q^{cf}_{jt_1}(\tau\,|\,\bar{z})\right] = \bar{z}^{\intercal}\left(\mathbb{E}\left[\gamma_{jt_1}(\tau) - \gamma^{cf}_{jt_1}(\tau)\right]\right),$$

with the term in parentheses being equivalent to the left hand side of (4.10), ultimately establishing the claimed second identity in (4.11).

It is furthermore insightful to frame the effect of the CBA coverage drop, as formalized in (4.11), as the portion of the actual temporal change that can be explained exclusively by the CBA coverage drop. To see this, define the factual temporal change of earnings for worker-type $\bar{z}$ at the τth quantile as

$$\Delta^{f}_{t_1,t_0}(\tau\,|\,\bar{z}) \equiv \mathbb{E}[Q_{jt_1}(\tau\,|\,\bar{z}) - Q_{jt_0}(\tau\,|\,\bar{z})], \quad (4.12)$$

and, analogously, the counterfactual temporal evolution as

$$\Delta^{cf}_{t_1,t_0}(\tau\,|\,\bar{z}) \equiv \mathbb{E}[Q^{cf}_{jt_1}(\tau\,|\,\bar{z}) - Q_{jt_0}(\tau\,|\,\bar{z})]. \quad (4.13)$$

Taking the difference of (4.12) and (4.13), it is immediately apparent that this is equivalent to the right hand side of (4.11), i.e., $\mathbb{E}[Q^{cf}_{jt_1}(\tau\,|\,\bar{z}) - Q_{jt_1}(\tau\,|\,\bar{z})]$, and hence represents the isolated effect of the CBA coverage drop:

$$\Delta^{f}_{t_1,t_0}(\tau\,|\,\bar{z}) - \Delta^{cf}_{t_1,t_0}(\tau\,|\,\bar{z}) = \bar{z}^{\intercal}\left(\left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}]\right)\mathbf{I}_K\right]\beta(\tau)\right).$$

Between-inequality measures

In the following, the notion of examining the effect of a CBA coverage drop on a between-inequality measure as proposed in Oka and Yamada (2023) is introduced. Following their remarks, the group-level aggregates for the factual between-group inequality measure is defined as

$$\Delta^{B}(\tau \mid \bar{z}_A, \bar{z}_B) \equiv \mathbb{E}\left[Q_{jt_1}(\tau \mid \bar{z}_B) - Q_{jt_1}(\tau \mid \bar{z}_A)\right].$$

Analogously, the counterfactual between-group inequality measure is given by

$$\Delta^{B,cf}(\tau \mid \bar{z}_A, \bar{z}_B) \equiv \mathbb{E}\left[Q^{cf}_{jt_1}(\tau \mid \bar{z}_B) - Q^{cf}_{jt_1}(\tau \mid \bar{z}_A)\right].$$

Intuitively, this between-inequality measure describes the extent to which, on average, the value of the τth conditional quantile for a group of workers with characteristics $\bar{z}_A$ is different from that of workers with characteristics $\bar{z}_B$.[11] It thus provides a measure that indicates the difference in earnings at a given relative position of the productivity distribution, arising from systematically different returns to worker characteristics.

Considering the difference between the factual and counterfactual measures of between-inequality, as shown in (4.11), the isolated effect of the decline in CBA coverage rates on the between-inequality measure can be expressed as:

$$\begin{aligned}\Delta^{B}(\tau \mid \bar{z}_A, \bar{z}_B) - \Delta^{B,cf}(\tau \mid \bar{z}_A, \bar{z}_B) &= \Delta^{f,cf}(\tau \mid \bar{z}_B) - \Delta^{f,cf}(\tau \mid \bar{z}_A) \\ &= (\bar{z}_B^{\intercal} - \bar{z}_A^{\intercal})\left(\left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}]\right)\mathbf{I}_K\right]\ \beta(\tau)\right). \qquad (4.14)\end{aligned}$$

Again, it is insightful to frame this isolated effect as the part of the between-inequality measure's temporal evolution that is explained by the CBA coverage drop alone. Define the factual and the counterfactual temporal evolutions of the between-inequality measure as

$$\Delta^{B,f}_{t_1,t_0}(\tau \mid \bar{z}_A, \bar{z}_B) = \Delta^{f}_{t_1,t_0}(\tau \mid \bar{z}_B) - \Delta^{f}_{t_1,t_0}(\tau \mid \bar{z}_A), \qquad (4.15)$$

$$\Delta^{B,cf}_{t_1,t_0}(\tau \mid \bar{z}_A, \bar{z}_B) = \Delta^{cf}_{t_1,t_0}(\tau \mid \bar{z}_B) - \Delta^{cf}_{t_1,t_0}(\tau \mid \bar{z}_A). \qquad (4.16)$$

It is then straightforward to show that the difference $\Delta^{B,f}_{t_1,t_0}(\tau \mid \bar{z}_A, \bar{z}_B) - \Delta^{B,cf}_{t_1,t_0}(\tau \mid \bar{z}_A, \bar{z}_B)$ is equal to the derived expression from (4.14) by the same token as above.

[11] Oka and Yamada (2023) provide an intuitive graphical intuition for this concept of between-inequality.

Within-inequality measures

Lastly, again following Oka and Yamada (2023), a within-group inequality measure is introduced. The within-group inequality measure assesses the spread of the distribution for a specified worker type $\bar{z}$ in terms of differences in quantiles. Specifically, to examine whether the CBA coverage drop led to a compression or widening of the within worker-type specific earnings distribution, the 85-15, 85-50, and 50-15 quantile differences in log earnings are considered below.[12]

Formally, the expected value for the factual within-inequality measure, which consists of comparing the τ_Ath and τ_Bth quantiles for a specific worker-type $\bar{z}$, reads as follows:

$$\Delta^{W}(\tau_A, \tau_B \,|\, \bar{z}) \equiv \mathbb{E}\left[Q_{jt_1}(\tau_B \,|\, \bar{z}) - Q_{jt_1}(\tau_A \,|\, \bar{z})\right].$$

Analogously, the counterfactual within-group inequality arises as

$$\Delta^{W,cf}(\tau_A, \tau_B \,|\, \bar{z}) \equiv \mathbb{E}\left[Q^{cf}_{jt_1}(\tau_B \,|\, \bar{z}) - Q^{cf}_{jt_1}(\tau_A \,|\, \bar{z})\right].$$

Similar to the between-inequality measure, the isolated effect of the drop in CBA coverage can be rephrased in terms of the quantile-specific difference for a given worker-type $\bar{z}$ as in (4.11), i.e.,

$$\begin{aligned}\Delta^{W}(\tau_A, \tau_B \,|\, \bar{z}) - \Delta^{W,cf}(\tau_A, \tau_B \,|\, \bar{z}) &= \Delta^{f,cf}(\tau_B) - \Delta^{f,cf}(\tau_A) \\ &= \bar{z}^{\intercal}\left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}]\right)\mathbf{I}_K\left(\beta(\tau_B) - \beta(\tau_A)\right)\right]. \qquad (4.17)\end{aligned}$$

The temporal and the counterfactual temporal evolution over time are easily derived for this case as well. Analogously to (4.15) and (4.16), they arise as

$$\Delta^{W,f}_{t_1,t_0}(\tau_A, \tau_B \,|\, \bar{z}) = \Delta^{f}_{t_1,t_0}(\tau_B \,|\, \bar{z}) - \Delta^{f}_{t_1,t_0}(\tau_A \,|\, \bar{z}), \qquad (4.18)$$

$$\Delta^{W,cf}_{t_1,t_0}(\tau_A, \tau_B \,|\, \bar{z}) = \Delta^{cf}_{t_1,t_0}(\tau_B \,|\, \bar{z}) - \Delta^{cf}_{t_1,t_0}(\tau_A \,|\, \bar{z}). \qquad (4.19)$$

Similarly to the cases above, the difference $\Delta^{W,cf}_{t_1,t_0} - \Delta^{W,f}_{t_1,t_0}$ equals the expression in (4.17).

[12] Considering these particular quantile differences follows the convention of using the 85th (15th) percentile instead of the 90th (10th) for the upper (lower) region of the distribution when using social security data from the IAB (compare, e.g., Dustmann et al., 2009).

4.5.3 Unconditional distributional effects of CBA coverage drop

The method presented above can also be used to derive the unconditional effect of de-unionization on the distribution of log daily earnings resulting from the implied shifts in the earnings structure. Earlier studies that considered the unconditional effect of de-unionization conceptually treated the phenomenon as a composition effect, separate from shifts in returns to worker observables, referred to as the 'wage structure' effect (e.g., Dustmann et al., 2009, 2014; Baumgarten et al., 2020; Biewen and Seckler, 2019). This stands in contrast to the notion underlying the grouped quantile regression approach presented above, which explicitly models the impact of a compositional shift on the structure of earnings, i.e., returns to worker observables. In the following, we outline an approach to assess the unconditional effect implied by the two-step model described above. This procedure allows us to examine the impact of de-unionization in Germany through the lens of shifts in the earnings structure, thereby deviating from earlier approaches and providing new evidence on its unconditional distributional effect.

To estimate the effect of the decline in CBA coverage between 1996 and 2014, the approach proposed by Machado and Mata (2005) is used. This involves the simulation of earnings observations from a conditional quantile regression model that characterizes the conditional distribution of earnings. To simulate earnings outcomes, we randomly select worker-level observables, z^s_{ijt}, and compute the corresponding earnings at a randomly chosen percentile, τ^s_i, of the conditional earnings distribution. Here, the superscript s refers to a specific simulation draw. Details of the simulation process are outlined in appendix C.4. Technically, this reads

$$\{w^s_{ijt} = \left(z^s_{ijt}\right)^\intercal \widehat{\gamma}_{jt}(\tau^s_i)\}^S_{s=1},$$

where group-specific parameter estimates, $\widehat{\gamma}_{jt}(\cdot)$, as described in (4.6), are used. S denotes the number of drawn simulation samples, which is 500 in this case. Since the group-level approach described above models the impact of the CBA coverage decline on the earnings structure, it is possible to simulate the log daily earnings for 2014 under the CBA coverage levels of 1996. Formally, using the implied vector of counterfactual returns as in (4.8), this is expressed as:

$$\{w_{ij,2014}^{s,cf} = \left(z_{ij,2014}^{s}\right)^{\intercal} \widehat{\gamma}_{j,2014}^{cf}(\tau_i^s)\}_{s=1}^{S}.$$

In this approach, the decline in CBA coverage from 1996 to 2014 exclusively affects the structure of earnings, suggesting that the counterfactual CBA coverage in 2014 implies potentially different returns. Hence, by using the factual and counterfactual distributions of log earnings, it is possible to obtain quantiles and quantile ratios to analyze the impact of the decline in CBA coverage on these distributional statistics. To that end, define $v(\cdot)$ as the unconditional distributional statistic, which refers to either the unconditional quantiles or the differences in unconditional quantiles of log earnings (specifically, the 85-15, 50-15, and 85-50 quantile differences). Using this, we can compute the following measures:

$$\text{Factual temp. evolution:} \quad \Delta_{2014} = v\left(f(w_{ij,2014}^{s})\right) - v\left(f(w_{ij,1996}^{s})\right), \tag{4.20}$$

$$\text{counterfactual temp. evolution:} \quad \Delta_{2014}^{cf} = v\left(f(w_{ij,2014}^{s,cf})\right) - v\left(f(w_{ij,1996}^{s})\right), \tag{4.21}$$

$$\text{isolated CBA effect:} \quad \Delta^{CBA} = \Delta_{2014} - \Delta_{2014}^{cf} = v\left(f(w_{ij,2014}^{s})\right) - v\left(f(w_{ij,2014}^{s,cf})\right). \tag{4.22}$$

4.6 Results

4.6.1 Impact of CBA coverage on the structure of earnings

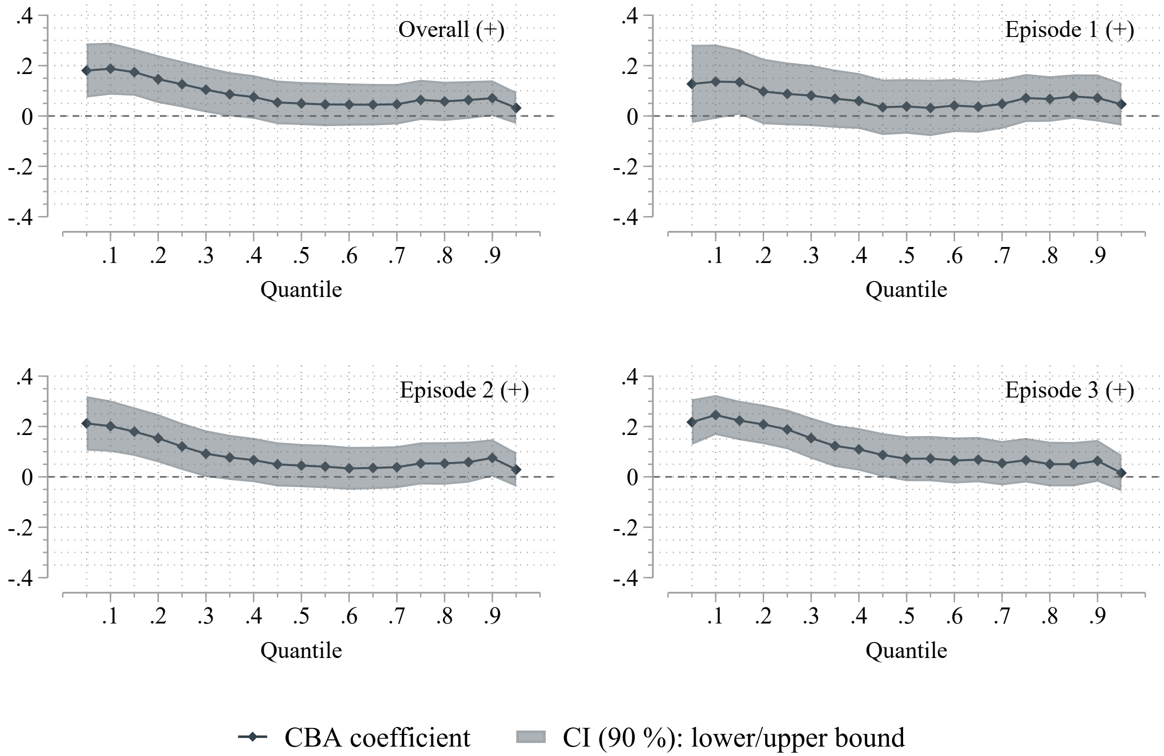
Figures 4.2 to 4.8 show the partial effect of the impact of CBA coverage on both the intercept as well as the returns to worker observables as specified in (4.4), incorporating potentially heterogeneous effects in different episodes. Additionally, an estimate of the overall effect, as implied by (4.2), is presented. These results indicate the extent to which changes in CBA coverage coincide with shifts in earnings differentials among worker types, providing a basis for examining the impact of a discrete decline in CBA coverage on measures of between- and within-group inequality.

Figure 4.2 depicts the effect of CBA coverage on the first-stage intercept. As only categorical worker-level covariates are considered in the first stage, the intercept refers to the base group of workers, i.e., male workers with vocational training aged 36-50. The results suggest that higher CBA coverage coincides with higher earnings at the lower to middle part of the distribution, reflecting a compression of earnings from below. This effect is essentially constant across the three episodes, with effect sizes being somewhat less pronounced in episode 1. Overall, these results are in line with previous findings that also highlighted that higher CBA coverage/union presence coincides with higher earnings growth at the lower tail of the male earning distributions (Dustmann et al., 2009, 2014). To understand the effect sizes, consider the impact at the 10th percentile in episode 3. The results suggest that a 10%-point reduction in CBA coverage coincides with a 2% decrease in real earnings.

Figures 4.3 and 4.4 illustrate how CBA coverage is associated with shifts in earnings differentials between workers with varying levels of educational attainment. As to the earnings penalty for workers with no formal education relative to those with vocational training (the base category), figure 4.3 suggests an equalizing effect of union-mediated wage setting during the most recent period (2008-2014). Specifically, a higher prevalence of CBA coverage generally coincides with less severe earning penalties for individuals with lower educational attainment compared to the base group at the middle to upper parts of the worker productivity distribution.[13] Conversely, in the earliest period considered (1996-2001), the effects indicate slightly detrimental impacts of higher CBA coverage at the upper part of the productivity distribution (see discussion of temporal heterogeneity below). As to the effect on the relative earnings surplus for university graduates, the results in figure 4.4 also suggest a temporally heterogeneous impact of CBA coverage. While there is consistent evidence of a negative effect of CBA coverage on the earnings surplus for university vs. vocational training at the very upper tail of the worker productivity distribution, the results are mixed at the lower to middle parts of the distribution. In the first episode, the negative effects are more pronounced, even in the middle part of the distribution, whereas findings for episode 3 indicate that higher CBA coverage is beneficial for workers with higher education at the lower to middle part of the distribution and decreases towards the upper quantiles.

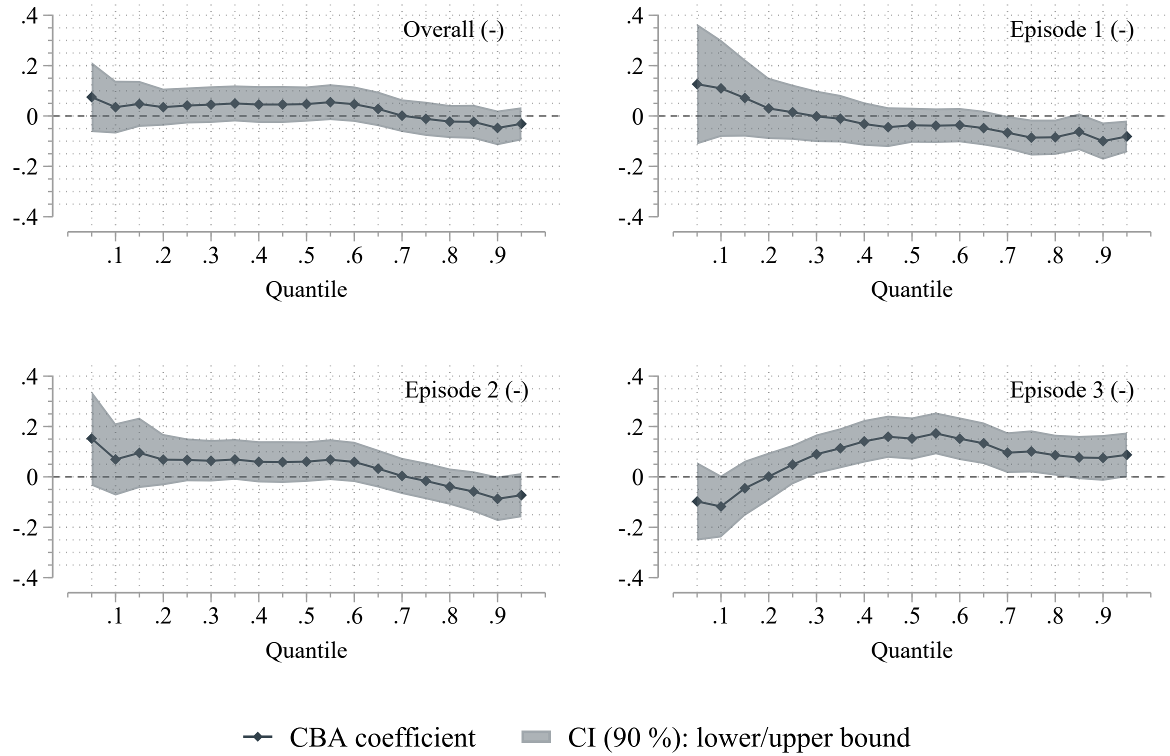
[13] Equalizing findings along the lines of educational attainments were also documented more recently for the US in Farber et al. (2021). Earlier contributions likewise indicated a tendency for smaller wage differentials under higher levels of centralization, especially among less-skilled educational groups within the 'unionized' (e.g., Card, 1996, and the references therein).

Figure 4.2: Effect of CBA coverage on the intercept (Base: male, vocational training, aged 36-50) across conditional quantiles



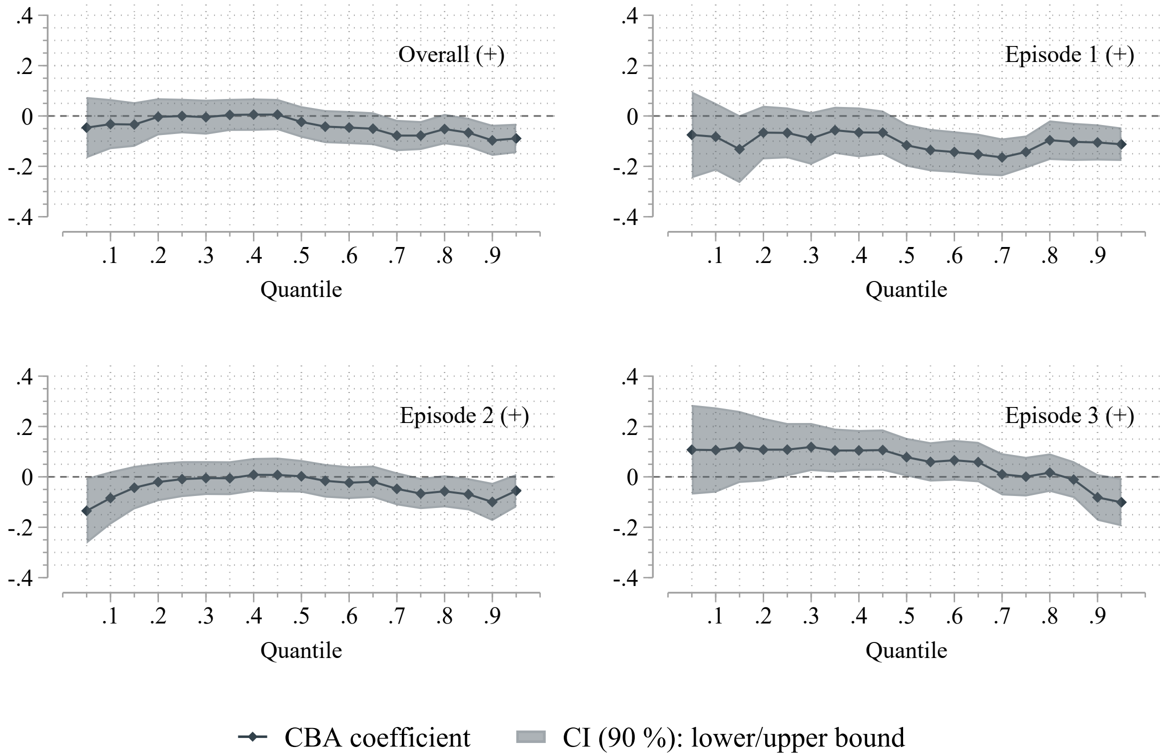
Notes: The figure illustrates the impact of CBA coverage on first stage returns (Intercept – refers to the base group of male workers with vocational training, aged 36-50) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The "Overall" effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.3: Effect of CBA coverage on return to no or other training across conditional quantiles



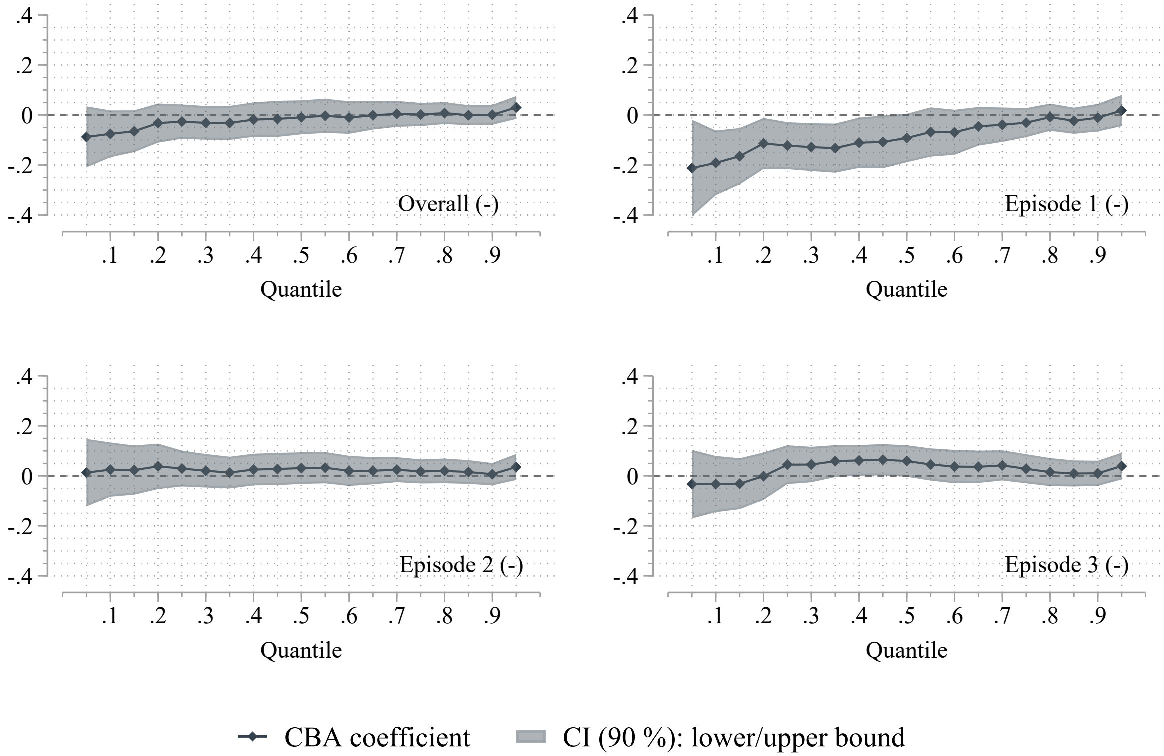
Notes: The figure illustrates the impact of CBA coverage on first stage returns to no formal education relative to vocational training (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The "Overall" effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.4: Effect of CBA coverage on return to higher education across conditional quantiles



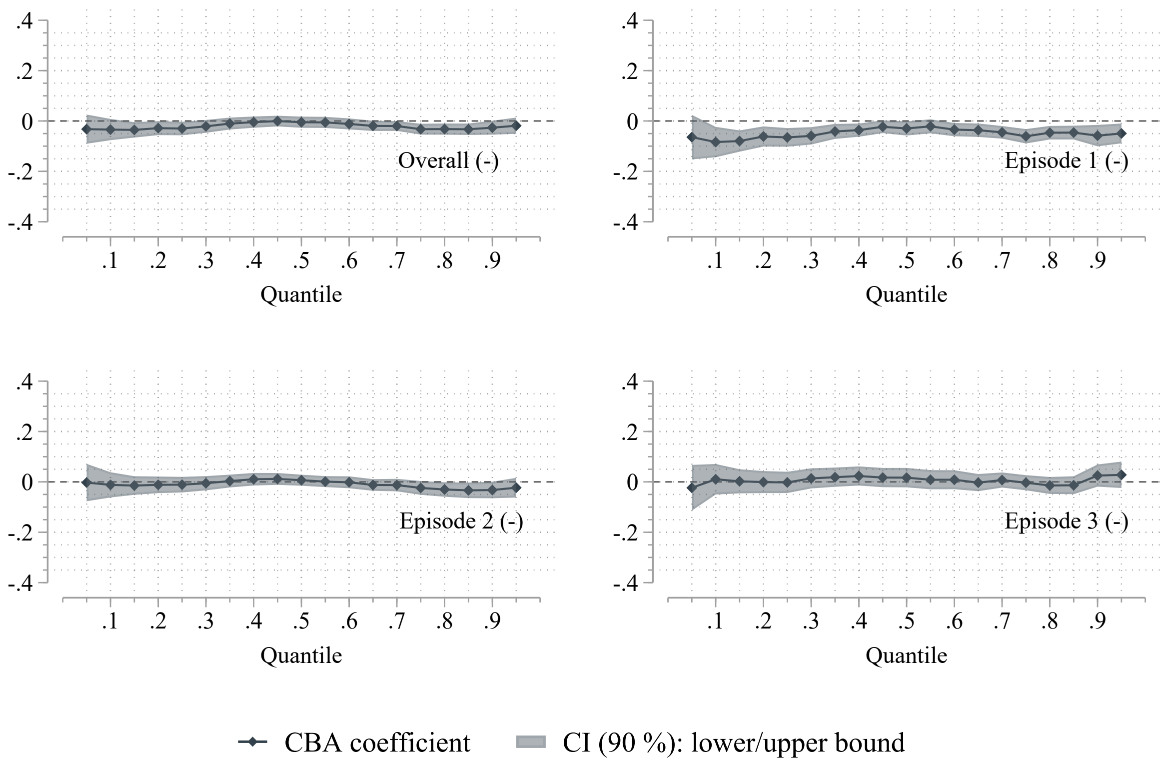
Notes: The figure illustrates the impact of CBA coverage on first stage returns to higher education relative to vocational training (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The “Overall” effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.5: Effect of CBA coverage on return to age group 18-25 across conditional quantiles



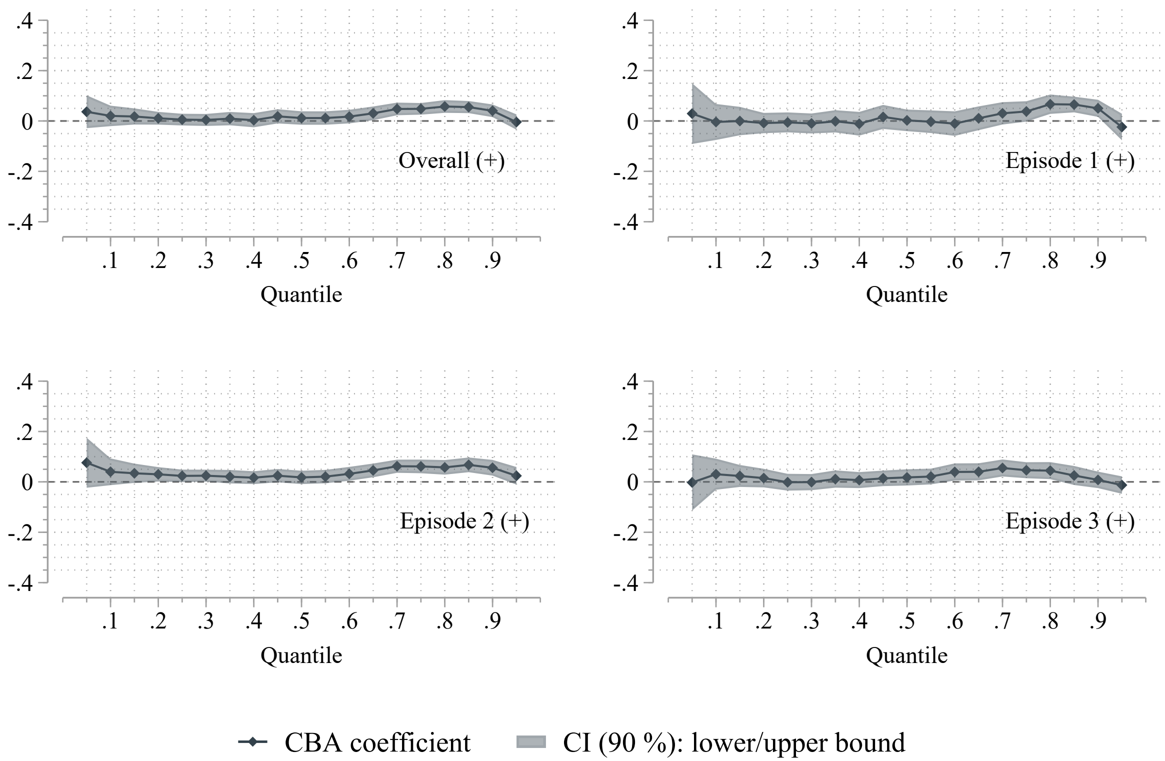
Notes: The figure illustrates the impact of CBA coverage on first stage returns to being in the age group 18-25 relative to being aged 36-50 (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The “Overall” effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.6: Effect of CBA coverage on return to age group 26-35 across conditional quantiles



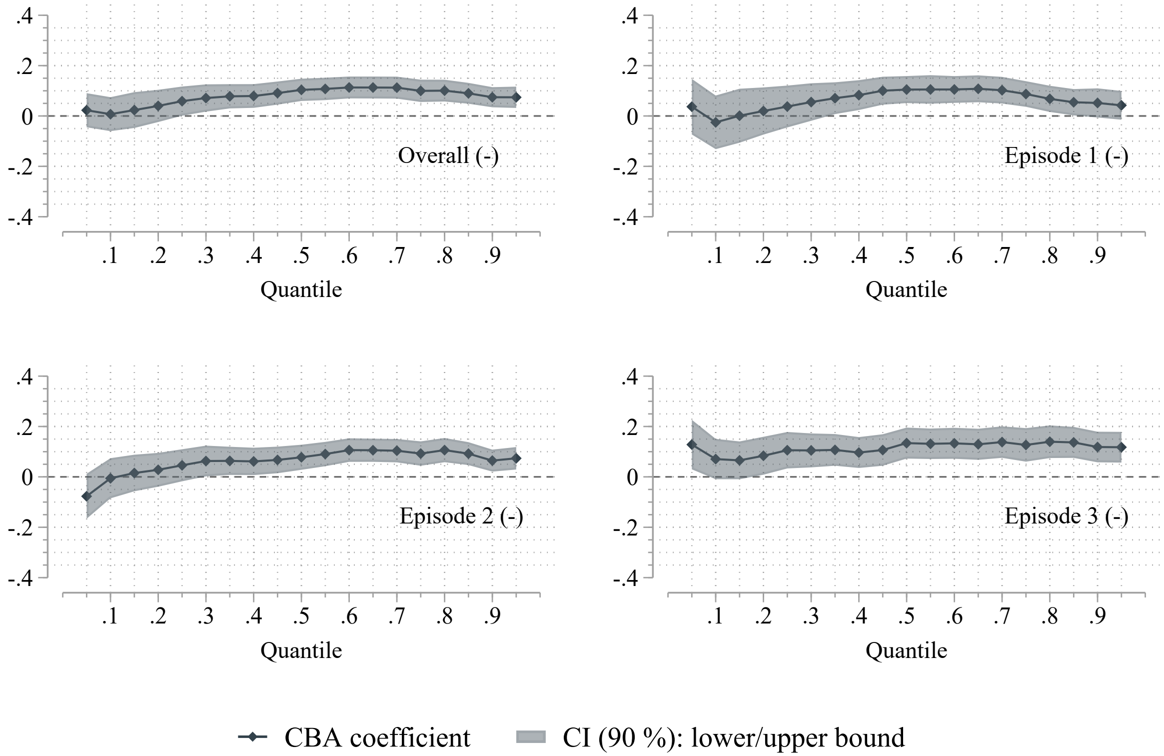
Notes: The figure illustrates the impact of CBA coverage on first stage returns to being in the age group 26-35 relative to being aged 36-50 (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The “Overall” effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.7: Effect of CBA coverage on return to age group 51-64 across conditional quantiles



Notes: The figure illustrates the impact of CBA coverage on first stage returns to being in the age group 51-64 relative to being aged 36-50 (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The "Overall" effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figure 4.8: Effect of CBA coverage on return to being female across conditional quantiles



Notes: The figure illustrates the impact of CBA coverage on first stage returns to being female vs. male (base group) for various conditional quantiles. The sign in parentheses indicates the direction of baseline effects (positive for wider, negative for narrower earnings differentials relative to the base group, compare table C.1). The effects, which vary by episode, correspond to the estimated returns of CBA coverage for specific periods (1996-2001, 2002-2007, and 2008-2014, respectively) as defined in (4.4). The "Overall" effect represents the coefficient for a model with a single CBA coverage variable, without episode differentiation, as per (4.2). The number of observations used to estimate the coefficients is quantile-dependent since each quantile represents a distinct second-stage regression. The minimum number of groups (sector × federal state × year) amounts to 2,479. The model includes sector, federal state, and year fixed effects, along with a linear time trend for each state and sector.
Source: LIAB QM2 9319, own calculations.

Figures 4.5, 4.6, and 4.7 illustrate how variations in CBA coverage prevalence coincide with shifts in earnings differentials across different potential experience groups. Overall, the findings suggest that CBA coverage prevalence only minimally affects these earnings differentials. However, one notable exception is the pronounced disadvantageous effect for young workers at lower quantiles of the productivity distribution in episode 1, as shown in figure 4.5. The only consistent result pertains to the age group 51-64, for which the findings suggest an additional benefit under more centralized wage settings relative to the base group at the upper tail of the worker productivity distributions, though the effect sizes are small.

Lastly, figure 4.8 displays the results on the female versus male earnings differential. The findings suggest that a more centralized wage-setting regime is associated with a smaller gender gap in earnings for full-time workers in the middle and upper parts of the productivity distribution. The effects are smaller and not (or only borderline) statistically significant at the lower end of the distribution. This pattern holds consistently across all three episodes examined and supports previous research indicating that greater centralization favors female workers.[14]

Temporal heterogeneity of the effect

Since some effects were found to be temporally heterogeneous, this could be interpreted as evidence that the effects of collective agreements depend on the general economic and labor market environment in which they are applied. The impact of collective agreements on earnings differentials between low- and medium-educated workers, as well as between younger and medium-aged workers, is found to have widened differences in real earnings during periods of economic uncertainty and labor market restructuring following German reunification (episode 1). During this period, unions experienced a decline in bargaining power. The loss of power was due to pressures induced by the fall of the Iron Curtain and the corresponding offshoring pressures, as well as spillover effects from employers who opted out of the system of collective negotiations in the new German federal states (Dauth

[14]A recent paper by Oberfichtner et al. (2020) finds no evidence of a mitigating effect of CBA on the within-firm gender gap. Using the same database as this study, they exploit discrete within-firm changes in CBA applicability while controlling for firm fixed effects. It should be noted, however, that such a firm-level analysis may suffer from the aforementioned problems in the German institutional setting. As also noted in Jäger et al. (2022, p. 69), "even an ideal firm-level experiment would leave open the question of equilibrium effects of sectoral bargaining through market spillovers or norms and expectations about pay (...)". The latter was the reason for using the group-level analysis, as described above.

et al., 2014; Jäger et al., 2022).

In particular, the fact that unions' bargaining power diminished might have prompted them to focus on their 'core clientele' in collective bargaining – with this core clientele being the base group of middle-aged male individuals with vocational training – somewhat neglecting workers who are not the 'median union member' (cf., Antonczyk et al., 2010). It is further noteworthy that the beneficial effect for female full-time workers was fairly constant over time, highlighting the important role that a more centralized system of wage setting seems to play for female workers.

4.6.2 Implications for between- and within-inequality

The following section discusses how the discrete drop in CBA coverage from 71% in 1996 to 52% in 2014 suggested shifts in measures for between- and within-earnings inequality. All empirical results presented below utilize the isolated effect of CBA coverage prevalence in 2014. That is, using the model formulated in (4.4), all effects that consider the isolated effect of CBA coverage prevalence take into account the effect in episode 3, i.e., $\widetilde{\beta}_k(\tau) \equiv \beta_{0k}(\tau) + \beta_{2k}(\tau)$.[15] To enhance readability, we abstracted from this in the formal introduction of the inequality measures previously presented in section 4.5.2.

CBA coverage drop and between-inequality

Figure 4.9 presents estimates for both the factual and counterfactual evolution of the between-group inequality measure as specified in (4.15) and (4.16). Here, $\bar{z}_A$ refers to the base group (male, vocational training, aged 36-50), and $\bar{z}_B$ corresponds to the respective worker group indicated in the figure. For the base group itself, the factual and counterfactual evolution of earnings between 1996 and 2014 is depicted as a reference. Additionally, figure 4.10 displays the difference between the factual and counterfactual effects alongside the 90% confidence interval.

[15] Formally, $\widetilde{\beta}_k(\tau)$ is the implied effect in episode 3 from (4.4) and corresponds to

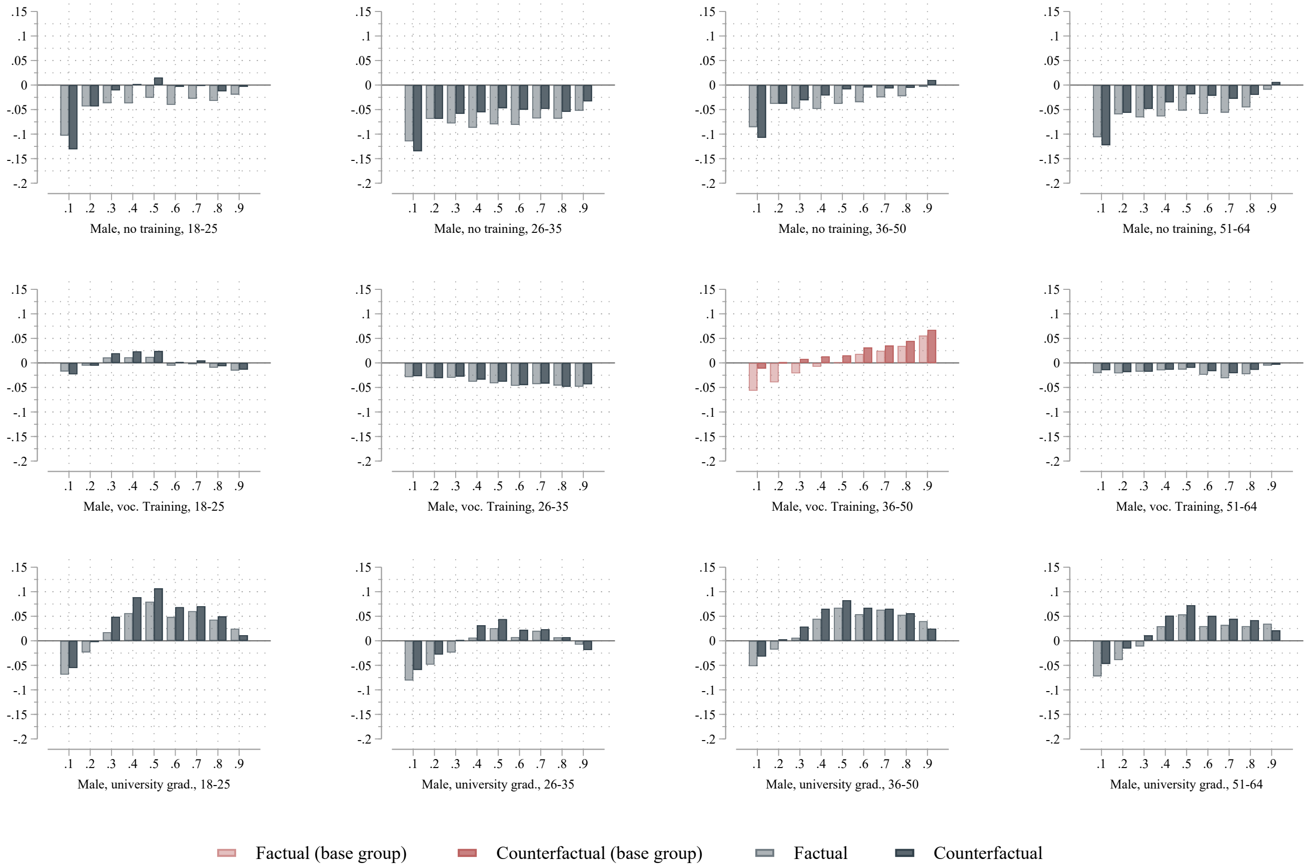
$$\left.\frac{\partial \gamma_{k,jt}(\tau)}{\partial u_{jt}}\right|_{D_{2t}=1} = \beta_{0k}(\tau) + \beta_{2k}(\tau) \equiv \widetilde{\beta}_k(\tau).$$

The results for the base group of workers imply factual losses in real earnings at the distribution's lower tail (-5% at the 10th percentile) and gains at the upper part (+5% at the 90th percentile). This is in line with previous contributions that have also documented a decline in real wages and earnings at the lower tail of the distribution and substantial gains at the top, which have been shown to fuel male earnings inequality (e.g., Dustmann et al., 2009). Had the CBA coverage not dropped, however, the results imply that the earnings losses for the base group at the lower tail would have actually been close to 0% or even slightly positive, whereas no statistically significant effects are detectable at the upper tail. Overall, these results imply that the drop in CBA coverage can explain the entire drop in earnings at lower quantiles, with effects being statistically significant for quantiles up to the median, as shown in figure 4.10.

The first row of figure 4.9 displays the temporal evolution of the earnings differential between the base group and male workers with no formal or other educational attainment for various age groups. The general pattern shows that the gap between low- and medium-educated worker types has factually increased over time. Specifically, the earnings penalty for workers with no formal education compared to those with vocational training has become more pronounced, highlighting tendencies previously attributed to cohort effects (see, e.g., Antonczyk et al., 2018). Again, the drop in CBA coverage can explain a substantial amount of the observed increase in this earnings wedge, accounting for the entire drop at the middle of the distribution for the age groups 18-25 and 36-50. The results in the first row of figure 4.10 further show that effects around the middle parts of the productivity distribution are statistically significant, whereas there is little to no effect heterogeneity along the dimension of age groups.

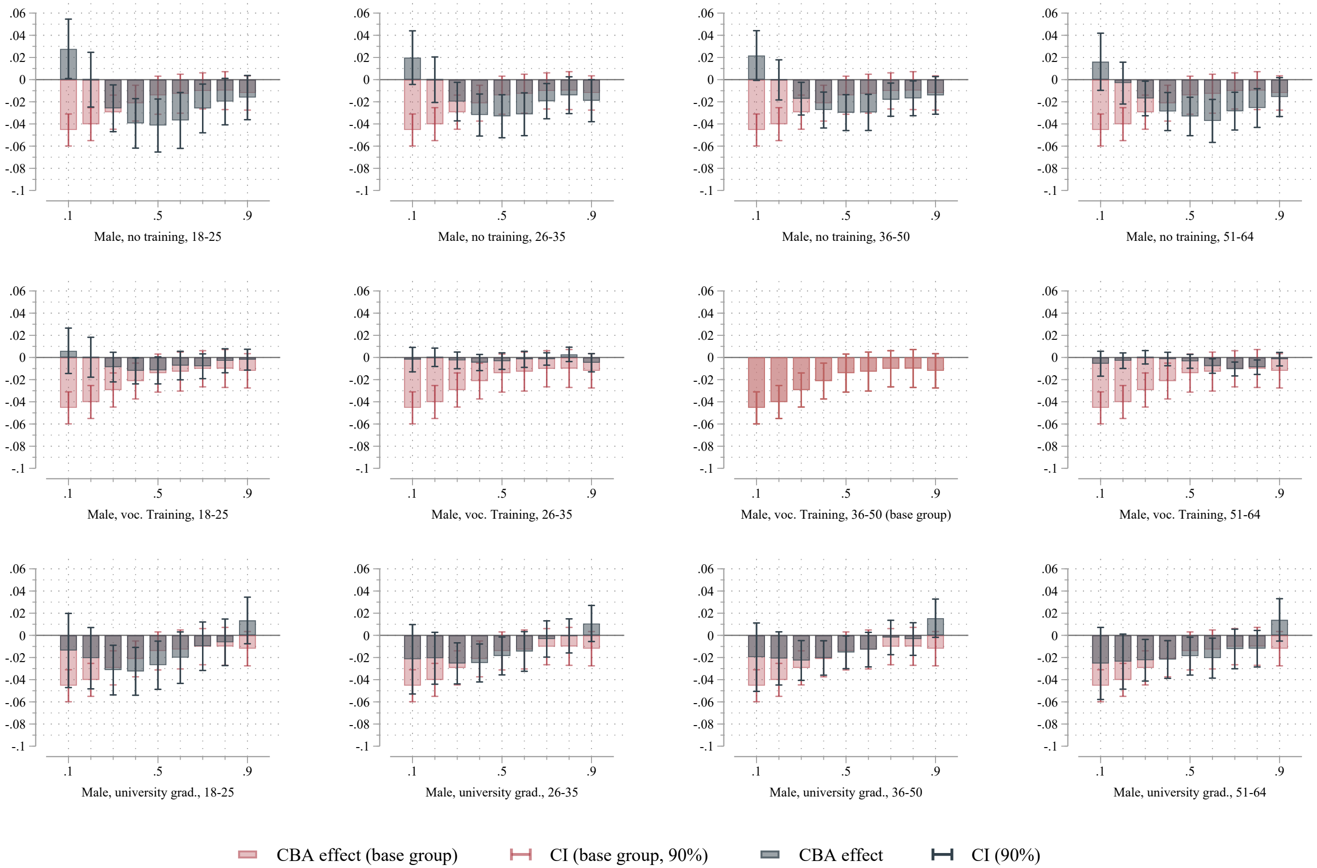
This lack of effect heterogeneity along age groups is corroborated by findings displayed in the second row of figure 4.9, showing that there are no large changes – neither factually nor counterfactually – in differentials between different age groups for otherwise similar-looking worker-types. As apparent from figure 4.10, effects are furthermore only borderline significant for the age groups 18-25 and 51-64 around the median and upper parts of the distribution, respectively, with effect sizes being barely economically relevant.

Figure 4.9: Factual and counterfactual evolution of between-inequality over time for various worker types (males)



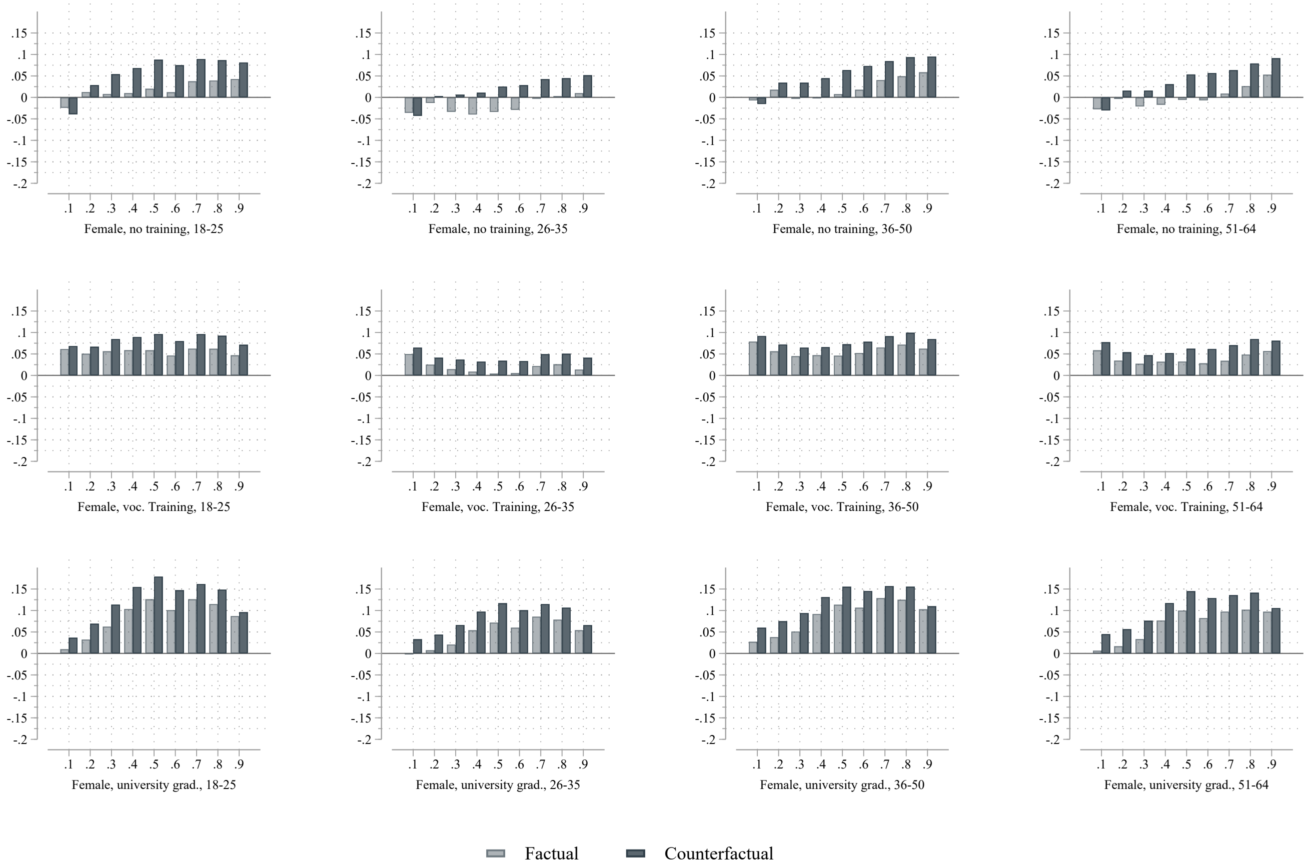
Notes: The figure shows the average factual and counterfactual evolution (1996 vs. 2014) of earnings relative to the base group (male, vocational training, aged 36-50) for various quantiles of the worker productivity distribution. Factual effects refer to an estimate of (4.15) with the discrete temporal comparisons being 2014 vs. 1996. The effects correspond to the comparison of effects that arise by considering $\bar{z}_B$ in (4.15), relative to the base group, $\bar{z}_A$. Hence, for the base group, the (counter)factual changes in earnings, rather than relative differences, are displayed in red. The counterfactual bars refer to a scenario in which the CBA coverage remained at its 1996 level, i.e., they reflect the factual evolution minus the effect of CBA coverage on the respective (combination of) coefficient(s) as given in (4.16). The current figure shows effects for male worker types.
Source: LIAB QM2 9319, own calculations.

Figure 4.10: Effect of CBA coverage on earnings for different worker types (males) and various quantiles



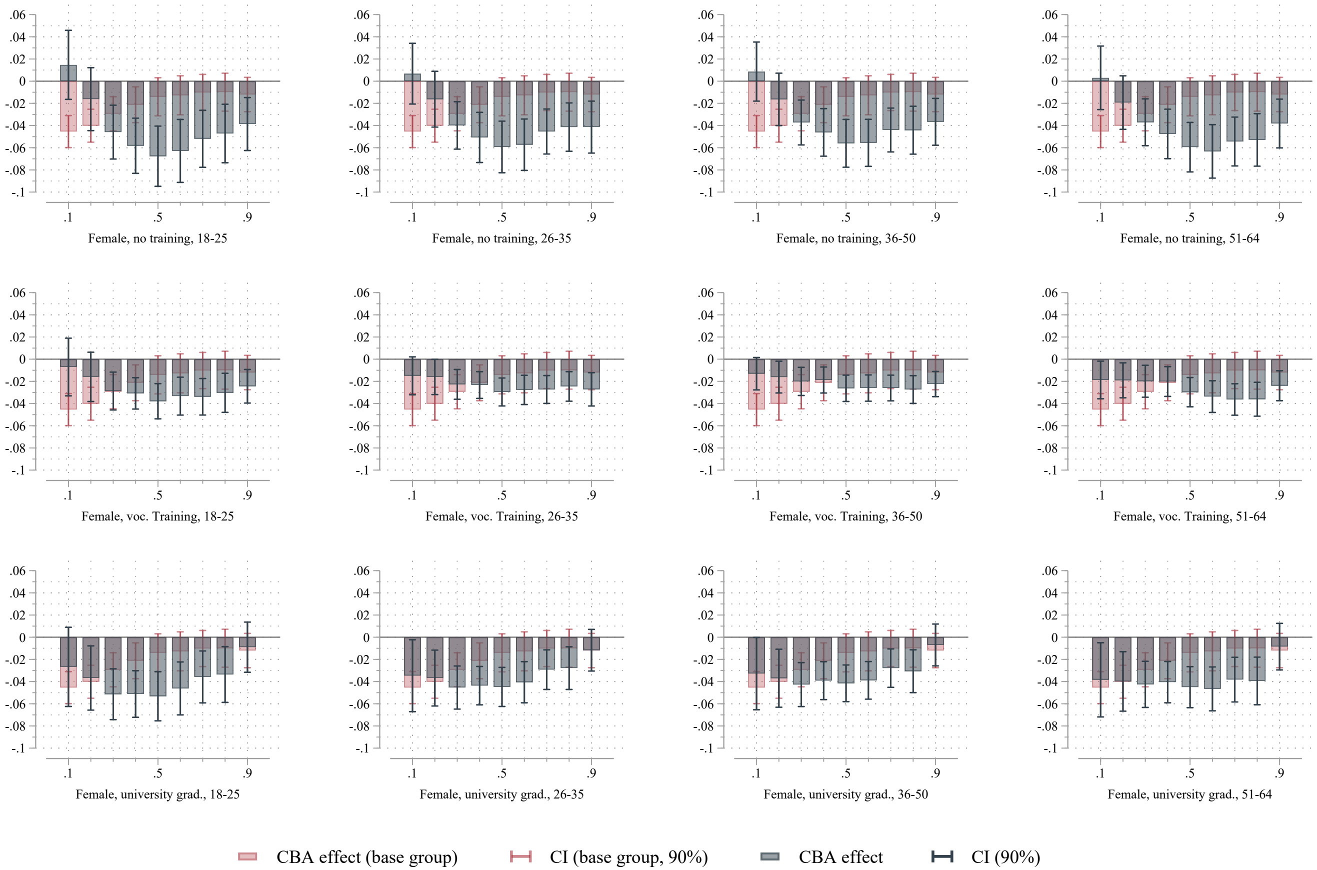
Notes: The figure illustrates the impact of the drop in CBA coverage from 1996 to 2014 on the measure for between-group inequality, i.e., the effect on earnings for various worker types relative to the base group along the distribution of log earnings. The effects refer to coefficient estimates from (4.14), with $\bar{z}_A$ being the base group (male, vocational training, aged 36-50). The current figure shows effects for male worker types. The base group is depicted in red and represents the effects on the intercept. All other effects refer to the (combined) effect of the CBA coverage drop for the respective worker group *relative* to the base group.
Source: LIAB QM2 9319, own calculations.

Figure 4.11: Factual and counterfactual evolution of between-inequality over time for various worker types (females)



Notes: The figure shows the average factual and counterfactual evolution (1996 vs. 2014) of earnings relative to the base group (male, vocational training, aged 36-50) for various quantiles of the worker productivity distribution. Factual effects refer to an estimate of (4.15) with the discrete temporal comparisons being 2014 vs. 1996. The effects correspond to the comparison of effects that arise by considering $\bar{z}_B$ in (4.15), relative to the base group, $\bar{z}_A$. Hence, for the base group, the (counter)factual changes in earnings, rather than relative differences, are displayed in red. The counterfactual bars refer to a scenario in which the CBA coverage remained at its 1996 level, i.e., they reflect the factual evolution minus the effect of CBA coverage on the respective (combination of) coefficient(s) as given in (4.16). The current figure shows effects for female worker types.
Source: LIAB QM2 9319, own calculations.

Figure 4.12: Effect of CBA coverage on earnings for different worker types (females) and various quantiles



Notes: The figure illustrates the impact of the drop in CBA coverage from 1996 to 2014 on the measure for between-group inequality, i.e., the effect on earnings for various worker types relative to the base group along the distribution of log earnings. The effects refer to coefficient estimates from (4.14), with $\bar{z}_A$ being the base group (male, vocational training, aged 36-50). The current figure shows effects for female worker types. The base group is depicted in red and represents the effects on the intercept. All other effects refer to the (combined) effect of the CBA coverage drop for the respective worker group *relative* to the base group.
Source: LIAB QM2 9319, own calculations.

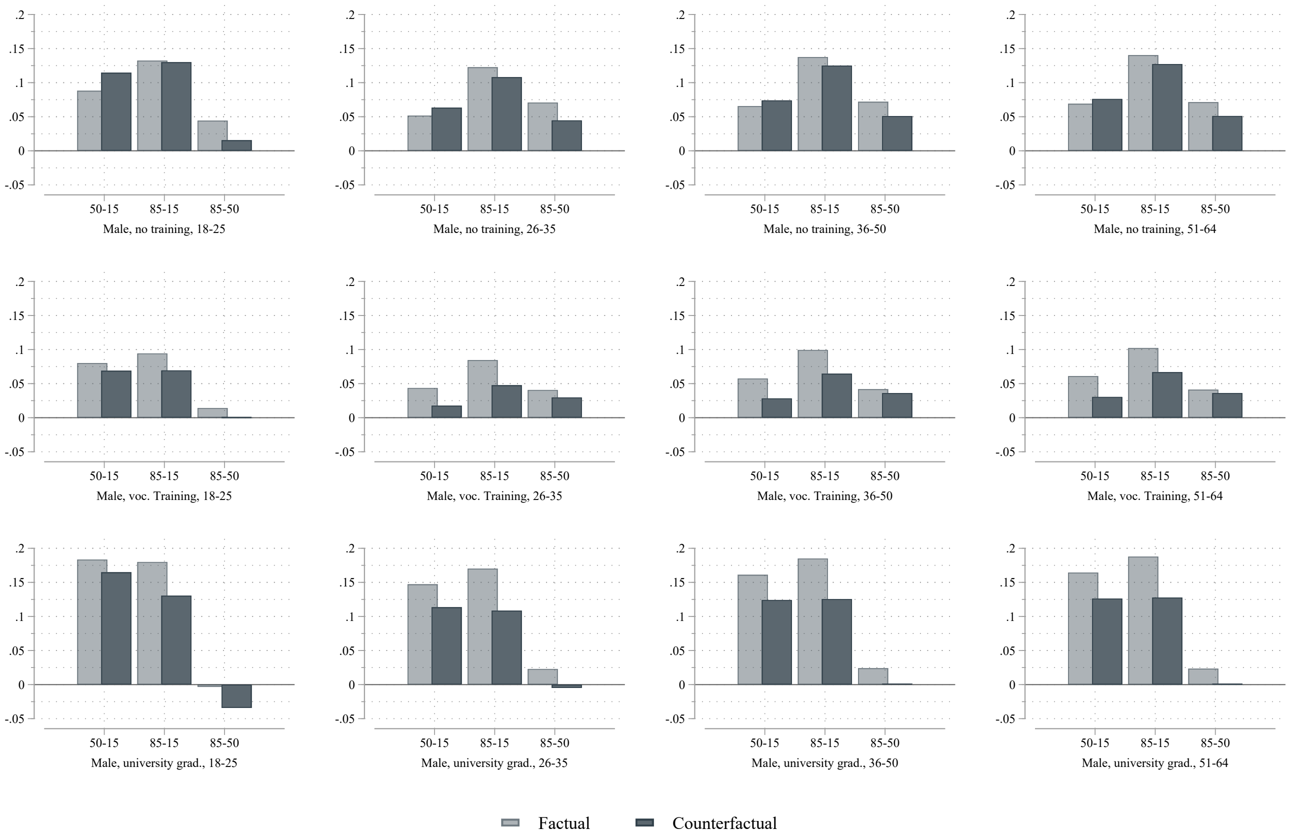
Lastly, the third row of figure 4.9 displays the temporal evolution of earnings differentials between workers with vocational and university training. The results highlight that earnings benefits for university graduates increased relative to workers with vocational training, especially at the middle to upper parts of the productivity distribution. In the absence of the CBA coverage drop, the increase in earnings differentials would have been even higher, but only at the lower to middle parts of the productivity distribution. As apparent from figure 4.10, effects are only found to be statistically significant for lower to middle quantiles. Again, there is virtually no effect heterogeneity across age groups.

Figure 4.11 displays the temporal evolution of earnings differentials for the female worker-types. Figure 4.12 shows the effects' statistical significance. As shown in figure 4.11, the factual relative increases in earnings for most female worker-types imply that the gender gap in earnings narrowed between 1996 and 2014 for virtually all worker-types. However, as indicated by the larger counterfactual bars, the decline in CBA coverage has largely attenuated this trend. To grasp the magnitude of the effect, consider, for example, the findings in figure 4.12 for young female workers with no formal education, where the largest effect is observed. Considering the median of the productivity distribution, the results for this worker-type imply that the decline in CBA coverage coincided with a statistically significant smaller relative increase in real earnings of about 7%-points (2% relative increase in earnings factually vs. 9% counterfactually). While the effect sizes vary for other types of female workers, the direction and pattern of the effects are similar, with the strongest effects observed in the middle to upper parts of the productivity distribution.

CBA coverage drop and within-inequality

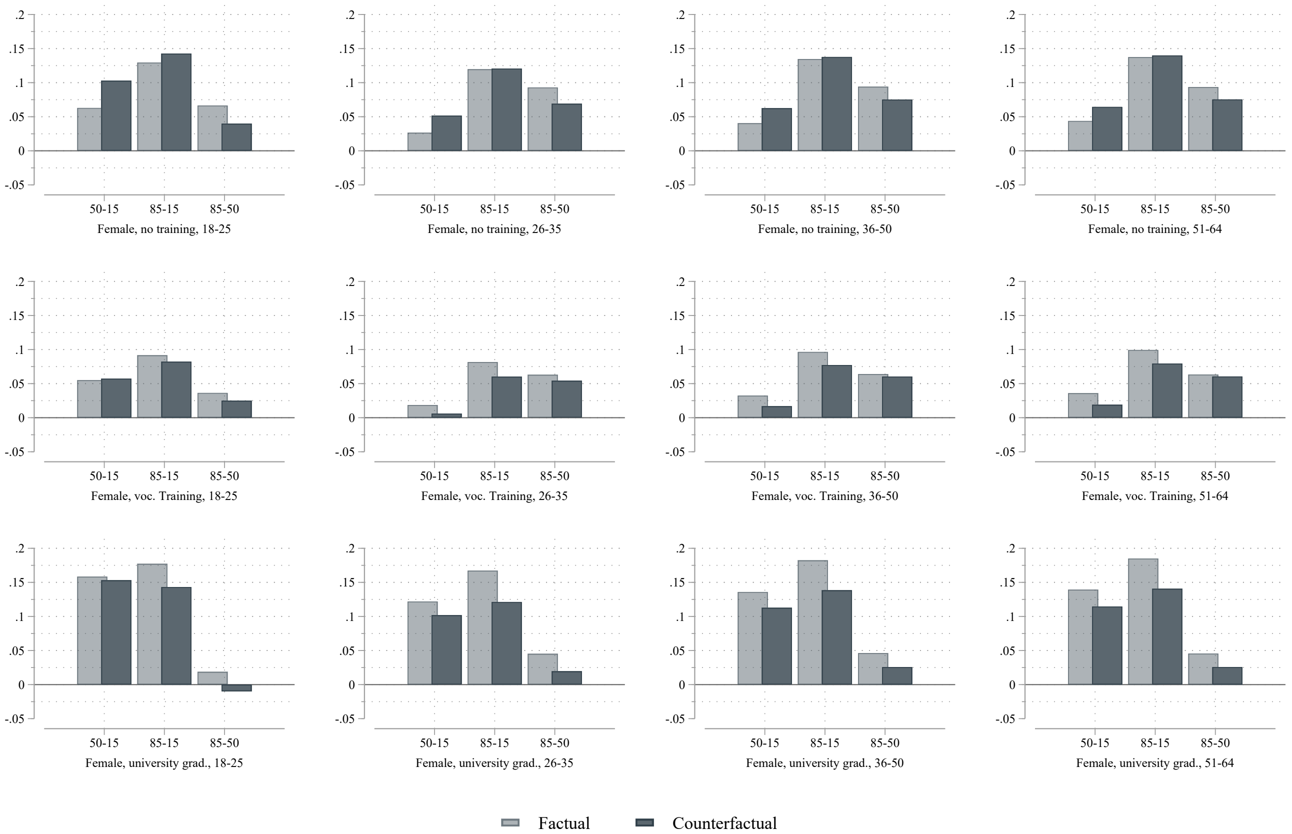
Figures 4.13 and 4.14 show estimates for the factual and counterfactual temporal evolution of the within-inequality measures for all considered worker-types as formalized in (4.18) and (4.19). Figures 4.15 and 4.16 show the statistical significance of the effects.

Figure 4.13: Factual and counterfactual evolution of earnings over time for various worker types (males)



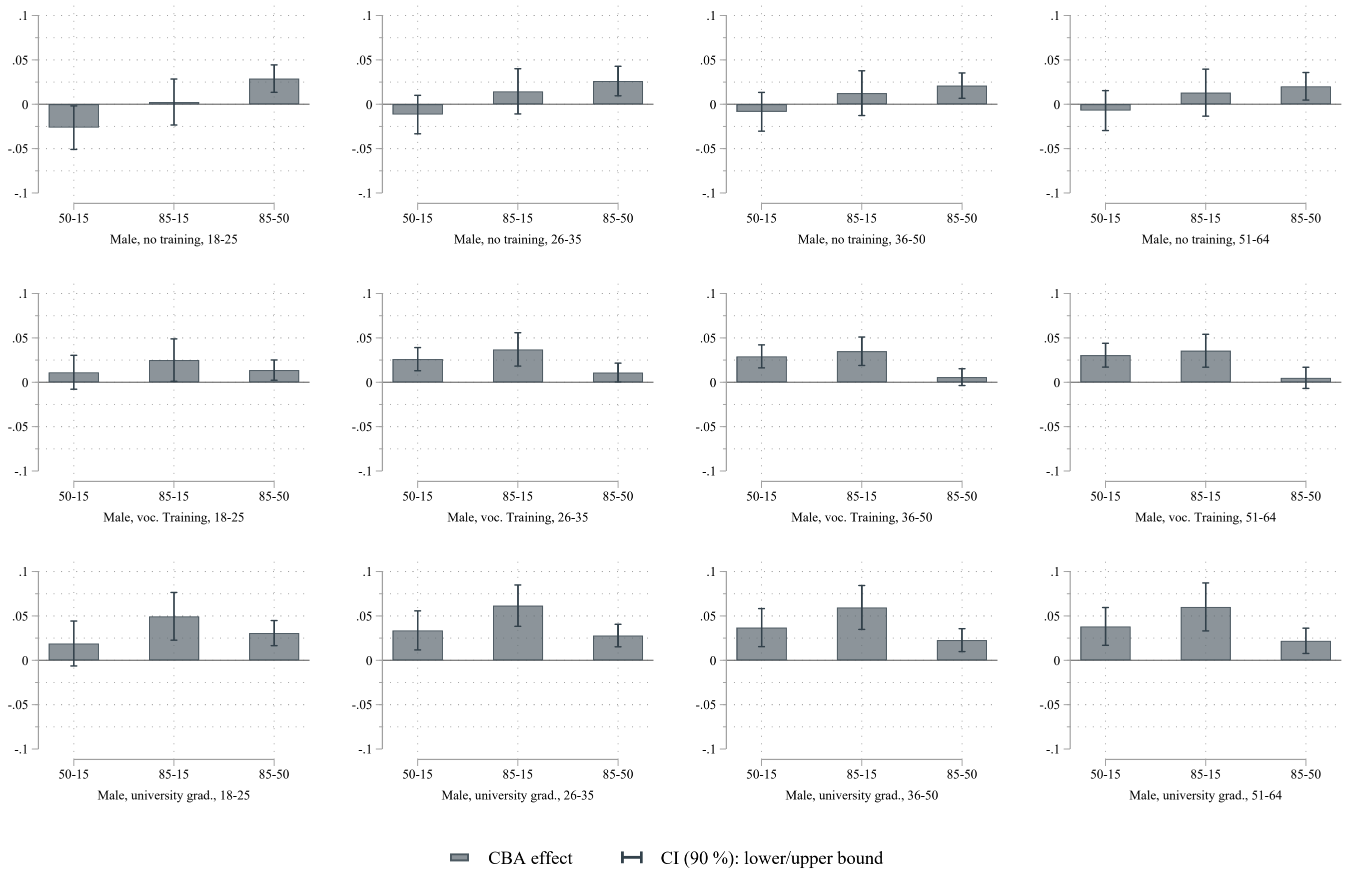
Notes: The figure shows both the factual and counterfactual changes in differences in quantiles of log earnings (ratios in levels of earnings) for various worker types, with differences referring to the comparison of the 50% vs. 15%, 85% vs. 50%, and 85% vs. 50% quantiles, respectively. Factual effects refer to discrete temporal comparisons (2014 vs. 1996) of the respective coefficients that define the groups as specified in (4.18). Correspondingly, the group "Male, vocational training, 36-50" refers to the temporal evolution of the constant coefficient. All other groups arise by subsequently adding the respective coefficients governed by $\bar{z}$. The counterfactual bars refer to a scenario in which the CBA coverage remained at its 1996 level, as formalized in (4.19). That is, they refer to the factual evolution minus the effect of CBA coverage on the respective (combination of) coefficient(s). The row dimension refers to effects along a given educational attainment (no training, vocational training, university graduate). The column dimension refers to effects along the four age groups (18-25, 26-35, 36-50, and 51-64). The current figure shows the effect for male worker types.
Source: LIAB QM2 9319, own calculations.

Figure 4.14: Factual and counterfactual evolution of earnings over time for various worker types (females)



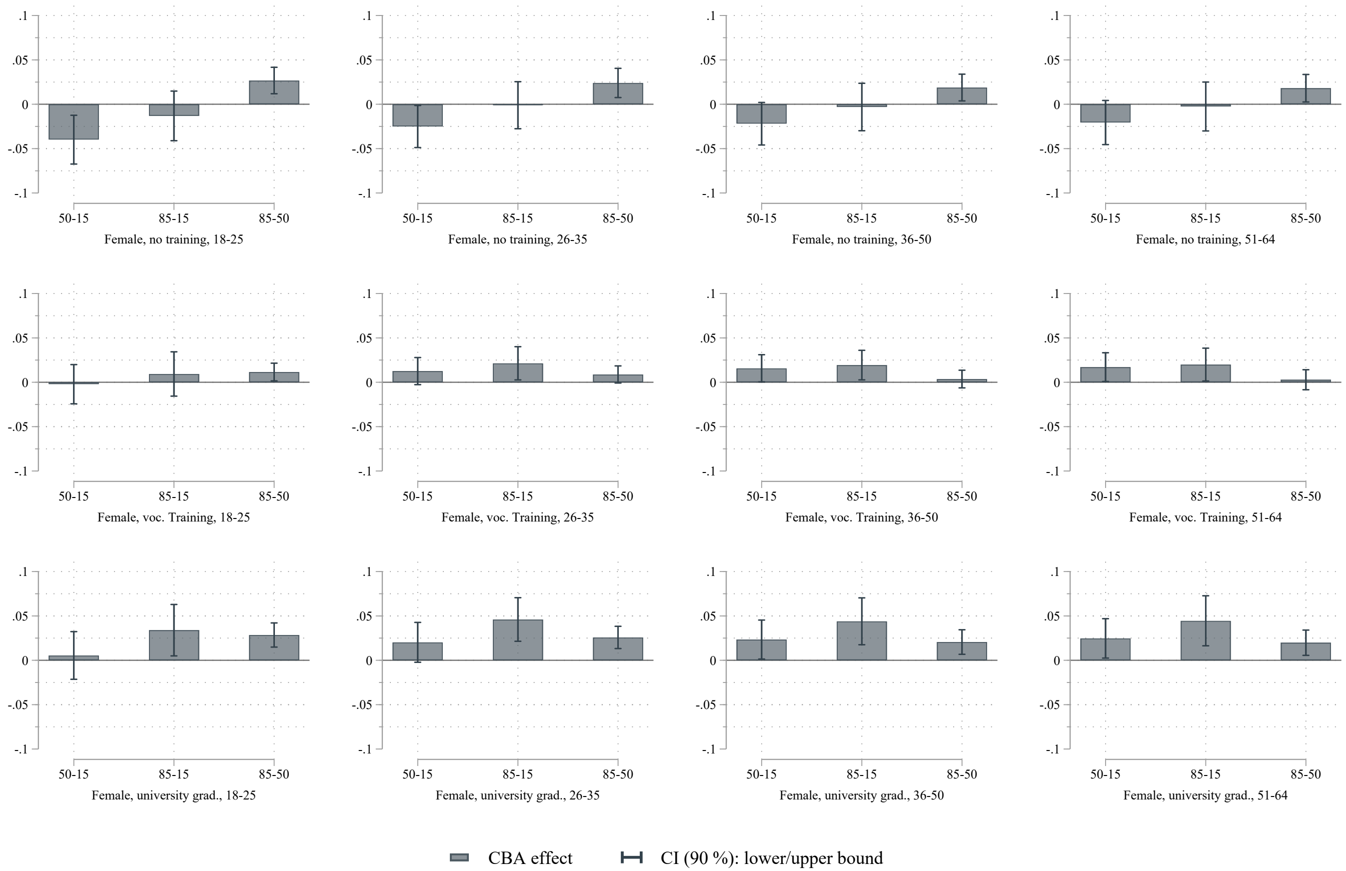
Notes: The figure shows both the factual and counterfactual changes in differences in quantiles of log earnings (ratios in levels of earnings) for various worker types, with differences referring to the comparison of the 50% vs. 15%, 85% vs. 50%, and 85% vs. 50% quantiles, respectively. Factual effects refer to discrete temporal comparisons (2014 vs. 1996) of the respective coefficients that define the groups as specified in (4.18). Correspondingly, the group "Male, vocational training, 36-50" refers to the temporal evolution of the constant coefficient. All other groups arise by subsequently adding the respective coefficients governed by $\bar{z}$. The counterfactual bars refer to a scenario in which the CBA coverage remained at its 1996 level, as formalized in (4.19). That is, they refer to the factual evolution minus the effect of CBA coverage on the respective (combination of) coefficient(s). The row dimension refers to effects along a given educational attainment (no training, vocational training, university graduate). The column dimension refers to effects along the four age groups (18-25, 26-35, 36-50, and 51-64). The current figure shows the effect for female worker types.
Source: LIAB QM2 9319, own calculations.

Figure 4.15: Effect of CBA coverage drop on quantile ratios within various worker types (males)



Notes: The figure shows the estimated effect alongside the 90% confidence interval of the drop in CBA coverage from 1996 to 2014 on differences in quantiles of log earnings (quantile ratios in level earnings) for various worker types, as specified in (4.17), with $\bar{z}$ determining the worker types shown in the figures and with differences referring to the comparison of the 50% vs. 15%, 85% vs. 50%, and 85% vs. 50% quantiles, respectively. The row dimension refers to effects along a given educational attainment (no training/vocational training/university graduate). The column dimension refers to effects along the four age groups (18-25, 26-35, 36-50, and 51-64). The current figure shows the effect for male worker types.
Source: LIAB QM2 9319, own calculations.

Figure 4.16: Effect of CBA coverage drop on quantile ratios within various worker types (females)



Notes: The figure shows the estimated effect alongside the 90% confidence interval of the drop in CBA coverage from 1996 to 2014 on differences in quantiles of log earnings (quantile ratios in level earnings) for various worker types, as specified in (4.17), with $\bar{z}$ determining the worker types shown in the figures and with differences referring to the comparison of the 50% vs. 15%, 85% vs. 50%, and 85% vs. 50% quantiles, respectively. The row dimension refers to effects along a given educational attainment (no training/vocational training/university graduate). The column dimension refers to effects along the four age groups (18-25, 26-35, 36-50, and 51-64). The current figure shows the effect for female worker types.
Source: LIAB QM2 9319, own calculations.

Similar to previous findings for the case of Germany, within-inequality, as measured by quantile ratios, increased across the board during the period under study.[16] For male worker types, the results in figure 4.13 show that inequality, measured by the 85-15 quantile difference in log daily earnings, increased substantially between 1996 and 2014. While this measure of overall inequality increased substantially for all worker types, there are differences among worker types based on educational attainment concerning the evolution of lower- and upper-tail inequality. For worker types with low and medium educational attainment (first and second row in figure 4.13), increases in both lower- and upper-tail inequality, indicated by the 50-15 and 85-50 quantile differences, rose comparably. Conversely, for worker-types with high educational attainment (third row in figure 4.13), upper-tail inequality remained stable, while lower-tail inequality increased considerably, with increases amounting to roughly 15% across all considered age groups.

Considering the counterfactual evolution in figure 4.13, it is apparent that virtually all of these increases would have been smaller in magnitude if CBA coverage had remained at its 1996 level, with effect sizes being strongest for worker-types with medium and higher educational attainment. By way of example, the findings in figures 4.13 and 4.15 suggest that overall inequality for male university graduates aged 26-35 would have been roughly 5%-points smaller if CBA coverage had not declined between 1996 and 2014, with the effect being significant at the 10% significance level. Overall, the results point to a compression effect of collective agreements, as indicated by the attenuated inequality increases in the counterfactual scenario.

Overall, the results for the female worker-types as presented in figures 4.14 and 4.16 are fairly comparable to the male worker-types. However, the magnitude of the effects – while pointing into the same direction – tend to be slightly smaller than for the case of male worker-types.

Taken together, these results suggest that in the past, a more centralized wage-setting regime compressed the structure of earnings towards (lower-)middle parts of the distribution. Specifically, for workers at the lower end of the unconditional earnings distribution (those with low educational attainment), the increasingly decentralized wage-setting regime led to an increase in upper-tail inequality, as measured by the 85-50 quantile ratio in real earnings. On the other hand, for workers who tend to be located at the middle to

[16]Compare, for example, Biewen and Seckler (2019) who considered a comparable period.

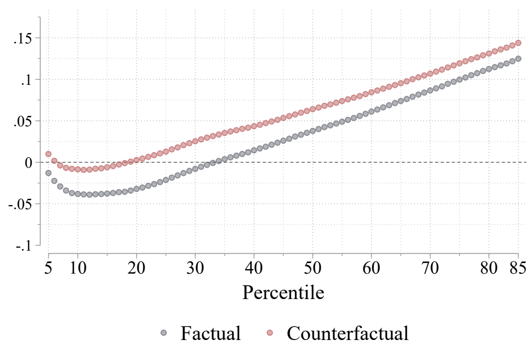
upper end of the unconditional earnings distribution (those with vocational or university training), inequality measures describing lower-tail inequality (50-15 quantile ratio) and overall inequality (85-15 quantile ratio) increased. Combined with the effects on the unconditional earnings distribution presented below, this aligns with earlier findings (e.g., Biewen and Seckler, 2019), adding a more nuanced characterization of the effect to the existing evidence.

4.6.3 Unconditional distributional effect

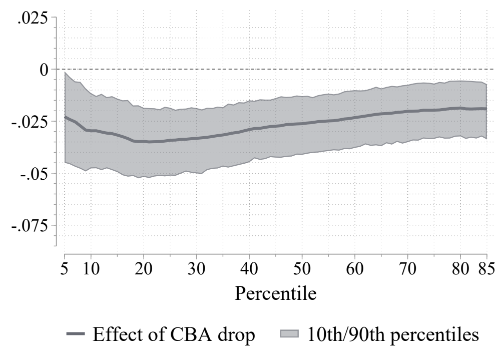
Based on the simulated sample of the unconditional distribution of log daily earnings, figure 4.17a illustrates the temporal evolution from 1996 to 2014, both factually and counterfactually, from the 5th to the 85th percentile of the unconditional earnings distribution, as formalized in (4.20) and (4.21).

Figure 4.17: Effect on unconditional distribution – temporal evolution and effect of CBA coverage drop (overall sample)

(a) Temporal change



(b) CBA effect



Notes: Subfigure 4.17a shows the log point change in factual and counterfactual earnings from 1996 to 2014 for the overall sample as given in equations (4.20) and (4.21) for several percentiles of the unconditional distribution of daily earnings. Subfigure 4.17b displays the part of the difference in daily earnings that can be ascribed to the drop in CBA coverage from 1996 to 2014, as given in (4.22). The isolated effects of the CBA coverage drop indicated by the solid line in subfigure 4.17b, correspond to the respective average over all simulated and bootstrapped worker-level observations for a given percentile value (see appendix C.4 for a detailed description of the procedure). The shaded area in subfigure 4.17b corresponds to the 10th and 90th percentiles of the sample of computed bootstrap values of the effect. These values thus represent the bounds of the 90% percentile bootstrap confidence intervals, as described in appendix C.4 and, e.g., in Firpo and Pinto (2016).
Source: Simulated data based on *LIAB QM2 9319*, own calculations.

Factually, real earnings have declined up to the 35th percentile, where this decrease in real daily earnings amounted to as much as 5%.[17] Conversely, however, if CBA coverage had remained at its 1996 level, the counterfactual evolution of earnings implies that declines would have been minuscule and would mostly indicate stagnation over time at lower percentiles. Moreover, the findings suggest that earnings would have already increased from the 20th percentile onward. The growth in earnings at upper percentiles is also found to be stronger under the counterfactual scenario, although the difference between the actual and counterfactual trajectories diminishes as one moves to the upper percentiles.

Put differently, as shown in figure 4.17b, the decline in CBA coverage between 1996 and 2014 accounts for a substantial fraction of the observed decline in real earnings, particularly at the lower tail of the distribution. By way of example, the decline in CBA coverage explains virtually all of the decline in earnings at the 20th percentile, which amounts to approximately 3.5%. Furthermore, figure 4.17b shows percentile bootstrap confidence bounds, indicating that the effect is significantly smaller than zero for all quantiles considered.[18] Overall, the induced shift in the structure of earnings is thus found to account for a substantial part of the observed changes along the unconditional distribution of real daily earnings.[19] These overall findings are consistent with previous contributions that considered a comparable period (Dustmann et al., 2014; Biewen and Seckler, 2019). Moreover, based on a methodological approach that explicitly models the impact of de-unionization through the implied shifts in the earnings structure, the findings presented in figure 4.17 provide additional evidence regarding the effect of de-unionization between 1996 and 2014 that goes beyond already existing studies. Unlike earlier contributions that considered only the aggregate and isolated effects of de-unionization, our findings reveal that de-unionization also influenced shifts in the earnings structure. This shift ultimately led to changes in the unconditional distribution of earnings.

[17]For a differential consideration of the male and female sub-populations, see figures C.2 and C.3 in appendix C.5. Note that the simulation-based results presented for the two sub-populations, as shown in figure C.2, are comparable to other studies that consider the evolution of earnings for both male and female sub-populations (Dustmann et al., 2009; Drechsel-Grau et al., 2022), suggesting consistency of the simulation-based approach.

[18]For a detailed explanation of the construction of these bounds using simulated data, see appendix C.4.

[19]Regarding gender differences, the results in figure C.2 show that while the shape of the effect is similar for both male and female sub-populations, the effect is larger for females. This reflects the beneficial impact of higher CBA coverage prevalence, as described in the worker-type findings above.

Table 4.2: Quantile differences in log daily earnings: evolution over time/effect of CBA coverage drop

	50-15			85-15			85-50		
	Overall	Male	Female	Overall	Male	Female	Overall	Male	Female
1996	0.380 [0.343,0.418]	0.323 [0.285,0.365]	0.434 [0.396,0.480]	0.717 [0.668,0.767]	0.664 [0.615,0.719]	0.766 [0.713,0.829]	0.337 [0.313,0.363]	0.341 [0.317,0.365]	0.332 [0.307,0.359]
2014 (fact.)	0.456 [0.421,0.488]	0.427 [0.388,0.467]	0.511 [0.481,0.540]	0.880 [0.831,0.930]	0.855 [0.804,0.909]	0.912 [0.868,0.956]	0.424 [0.398,0.449]	0.428 [0.402,0.452]	0.401 [0.369,0.433]
2014 (cf.)	0.450 [0.411,0.491]	0.422 [0.377,0.466]	0.501 [0.463,0.538]	0.867 [0.814,0.922]	0.844 [0.788,0.906]	0.894 [0.841,0.946]	0.417 [0.390,0.442]	0.422 [0.397,0.448]	0.393 [0.365,0.424]
Δ_t^{fact}	0.075 [0.033,0.116]	0.104 [0.057,0.146]	0.077 [0.029,0.127]	0.162 [0.112,0.213]	0.191 [0.139,0.237]	0.146 [0.086,0.201]	0.087 [0.059,0.116]	0.087 [0.059,0.118]	0.069 [0.038,0.100]
Δ_t^{cf}	0.070 [0.025,0.116]	0.099 [0.050,0.145]	0.066 [0.019,0.119]	0.150 [0.096,0.208]	0.180 [0.123,0.231]	0.127 [0.068,0.191]	0.080 [0.050,0.111]	0.081 [0.052,0.111]	0.061 [0.029,0.092]
Δ^{CBA}	0.005 [−0.007,0.016]	0.005 [−0.006,0.016]	0.011 [−0.010,0.028]	0.013 [−0.002,0.026]	0.011 [−0.003,0.024]	0.019 [−0.002,0.038]	0.007 [−0.001,0.016]	0.006 [−0.003,0.014]	0.008 [−0.002,0.020]

Notes: The table shows values and derived measures of unconditional quantile differences in log earnings based on the simulated sample of real log daily earnings, as described in detail in appendix C.4. The upper panel corresponds to the percentile differences based on the factual and counterfactual simulated values from 1996 and 2014. The lower panel displays results for the derived measures, including the factual and counterfactual difference, as well as the isolated effect of the CBA coverage drop, as given in equations (4.20), (4.21), and (4.22). The indicated values correspond to the respective average over all simulated and bootstrapped worker-level observations (see appendix C.4 for a detailed description of the procedure). The values in brackets correspond to the 10th and 90th percentiles of the computed bootstrap sample for the respective distributional statistic. These values, thus, represent the bounds of the 90% percentile bootstrap confidence intervals as described in appendix C.4 and, e.g., in Firpo and Pinto (2016).
Data: Simulated data based on *LIAB QM2 9319*, own calculations.

Table 4.2 shows the impact of the CBA coverage drop on unconditional quantile differences in log earnings. The full sample results reveal an increase in all assessed inequality measures from 1996 to 2014, aligning with prior research (e.g., Biewen and Seckler, 2019). Had CBA coverage stayed at its 1996 level, the rise in earnings inequality would have been less marked, with an economically relevant effect size. Specifically, the decrease in CBA coverage contributes approximately 7% to the observed reduction in the 50-15 quantile gap and 8% to the 85-15 and 85-50 quantile differences in log earnings.[20] However, based on the 90% percentile bootstrap confidence bounds, the isolated effect of the CBA coverage decline is not statistically significantly different from zero. Overall, the effect of de-unionization is found to be more pronounced in magnitude for female worker types across all measures considered, although these effects are not statistically significant either.

[20]Percentages refer to the ratio of the isolated CBA coverage effect to the actual change between 1996 and 2014, i.e., $\Delta^{CBA}/\Delta^{fact}$ (cf., table 4.2).

4.7 Conclusion

This paper analyzes the impact of the unprecedented de-unionization between 1996 and 2014 on the structure of earnings in Germany. To estimate the effect of the prevalence of collective agreements on earnings inequality, a grouped quantile regression approach (Chetverikov et al., 2016) is employed, which has not been applied to this research question before. This methodological choice is guided by the goal of aligning the empirical approach with the German institutional setting, as the latter complicates a firm- or employee-level analysis due to potential spillover effects to non-covered firms or employees. The method used here perfectly fits this level of treatment. In particular, unions, through the applicability of collective agreements, take effect at a comparatively high level of centralization in the German context. Since most collective agreements apply at the sector-region level rather than at the firm level, only the group-level variation in the CBA coverage prevalence is exploited in the empirical analysis. The approach used here reveals the effect's heterogeneity across various worker types as well as along the distribution of worker productivity. Furthermore, it easily accommodates the existing censoring in German social security data without the need for data imputation, by using a censored quantile regression approach (Chernozhukov and Hong, 2002; Chernozhukov et al., 2015).

Using a large-scale linked employer-employee database that combines administrative worker-level data with a firm-level survey (the *LIAB*), the results of the empirical analysis suggest that variations in the prevalence of CBA coverage affect the structure of earnings in Germany. This, in turn, has implications for measures of between- and within-inequality, i.e., the decline in CBA coverage that occurred between the mid-1990s and the mid-2010s is associated with distributional shifts. On the one hand, in line with earlier contributions (e.g., Dustmann et al., 2009, 2014), the group-level analysis suggests that a more centralized wage-setting regime coincides with a compression of earnings distribution from below for the 'core clientele' workers of unions (full-time, middle-aged, male workers with vocational training) but also shows detailed interactions of earnings returns with CBA coverage. This paints a more nuanced picture than earlier related contributions (Dustmann et al., 2009; Biewen and Seckler, 2019). Regarding earnings differentials relative to this reference group, higher CBA coverage is associated with a less negative penalty for full-time female workers, particularly in the middle to upper segments of the worker productivity distribution. Shifts in CBA coverage also affect earnings differentials between workers with different educational qualifications. Specifically, the effects point to a narrowing of the earnings gap at the lower end of the educational attainment groups, i.e., the penalty for workers without vocational training tends to be smaller relative to types of workers with vocational training in the middle and upper part of the productivity distribution if CBA coverage is higher. University graduates also benefit from a more centralized wage setting in times of economic booms, however, only at the lower to middle part of the productivity distribution, whereas detrimental effects were found at the very upper tail of the productivity distribution.

As to the effect of the observable decline in CBA coverage over time, the empirical analysis highlights the equalizing effect of a centralized wage setting system, which faced a pronounced decline in Germany over the period under consideration (1996-2014). It is demonstrated that the decline in CBA coverage rate in many cases accounted for a substantial fraction of the actually observable widening of earnings differentials, with the strongest effect being found along the lines of different educational attainments and gender. The empirical findings show that the gender gap in earnings between the considered female worker-types and the base group would have been substantially smaller if the CBA coverage had remained at its initial level. Regarding the empirical results for earnings differentials between differently educated worker-types, it was found that the observed increase in earnings differentials between workers with different levels of education is somewhat amplified by the decline in CBA coverage. That is, the earnings penalties for low- versus medium-educated worker types would have been much smaller if CBA coverage had not declined. One reading of the results could thus be that the estimated effects point to an additional ‘institutional’ layer that mediates changes in earnings and inequality due to, e.g., technology-augmented supply/demand shifts for certain worker-types.

Regarding the effect of de-unionization on within-inequality, the results are consistent with the ex ante intuition that a more centralized wage setting compresses earnings towards the median worker, which is the base group in the analysis. While the decrease in CBA coverage mainly affected upper-tail inequality (85-50 quantile ratio in earnings) for low-educated worker-types, it affected lower-tail inequality (50-15 quantile ratio) or overall inequality (85-15 quantile ratio) for medium- and high-educated worker-types, respectively. Effects for female workers were comparable but somewhat smaller in magnitude.

Lastly, by using an innovative combination of the group-level quantile regression approach with the simulation procedure of Machado and Mata (2005), it is demonstrated how the unconditional distributional effect of the CBA coverage drop can be obtained. Taking advantage of the fact that both the approach by Chetverikov et al. (2016) and the simulation approach proposed by Machado and Mata (2005) rely on conditional quantile regression models, a counterfactual earnings structure that incorporates the thought experiment of holding CBA coverage at its 1996 level can be used to simulate a counterfactual sample of log earnings in 2014. Ultimately, this allows us to obtain the isolated effect of the decline in CBA coverage along the unconditional distribution of earnings. Doing so, we find that the implied shifts in the structure of earnings translate into substantial shifts in the unconditional distribution of earnings, especially at the lower to lower-middle percentiles.

Overall, using a novel methodological approach to examine the effects of the unprecedented de-unionization in Germany, the results underscore the importance of collective agreements in Germany and provide a detailed characterization of the impact by demonstrating heterogeneities across different worker types. To conclude, the decline in union coverage in Germany not only resulted in larger earnings differentials between workers in covered and uncovered sectors but also highlighted important dimensions of heterogeneous effects by type of worker.

Appendix

C.1 Estimation of censored quantile regression coefficients (Chernozhukov and Hong, 2002)

The following outlines the three-step algorithm for estimating the parameters of a quantile regression model affected by top censoring, as described by Chernozhukov and Hong (2002).

Step 1 (Chernozhukov and Hong, 2002). In the first step, an initial rough set of suitable observations is formed, characterized by those observations that are statistically unlikely to be affected by censoring. To achieve this, define the indicator for *not* being censored as $\delta_{ijt} \equiv \mathbb{1}[w_{ijt} < c_{jt}]$, where c_{jt} is the threshold that applies in sector-region j at time t. Correspondingly, denote the true conditional propensity of *not* being censored as $\mathbb{P}[\delta_{ijt} = 1 \,|\, \bar{z}_{ijt}, c_{jt}] \equiv \bar{\tau}$, where $\bar{z}_{ijt}$ corresponds to a specific cell of worker-level covariates. This implies that, if the true quantile regression line were known and it were indeed linear in parameters, then the following would hold:

$$Q_{w_{ijt}}(\bar{\tau} \,|\, \bar{z}_{ijt}) = \bar{z}_{ijt}^{\intercal} \gamma_{jt}(\bar{\tau}) = c_{jt},$$

i.e., the $\bar{\tau}$th quantile of the conditional distribution given $\bar{z}_{ijt}$ would be exactly the censoring threshold (the border case). If one chose a covariate cell, $\check{z}_{ijt}$, such that $\mathbb{P}[\delta_{ijt} = 1 \,|\, \check{z}_{ijt}, c_{jt}] = \check{\tau}$, with $\check{\tau} > \bar{\tau}$, this would imply that $Q_{w_{ijt}}(\check{\tau} \,|\, \check{z}_{ijt}) < c_{jt}$, since, intuitively, a more ‘cautious’ covariate cell (in terms of the propensity of not being censored) would be considered.

Even though the true propensity of not being censored, $\mathbb{P}[\delta_{ijt} \,|\, z_{ijt}, c_{jt}]$, is generally unknown, it is possible to find a parametric (or non-parametric, cf. Oka and Yamada, 2023) estimate for the true propensity, i.e.,

$$p(\tilde{z}_{ijt}, c_{jt}) \equiv \hat{\mathbb{P}}[\delta_{ijt} = 1 \,|\, z_{ijt}, c_{jt}].$$

Following Chernozhukov and Hong (2002), $p(\cdot)$ is a probit link function and $\tilde{z}_{ijt}$ involves the worker-level characteristics and several interactions between them.[21] As noted by Chernozhukov and Hong (2002), the (almost sure) inconsistency of the estimate for the true propensity does not matter for the given purpose, as one only requires a suitable starting set of observations. That is, for a given τ, the set is defined such that it comprises those observations for which it approximately holds that $\mathbb{P}[\delta_{ijt} = 1 \,|\, z_{ijt}, c_{jt}] > \tau$, i.e.,

$$J_{0,jt}(k) = \{i : p(\tilde{z}_{ijt}, c_{jt}) > \tau + k\}, \tag{C.1}$$

with k being strictly between 0 and $1-\tau$. The latter can be thought of as compensating potential misspecifications of $p(\cdot)$ by defining a more cautious set of observations that is to be used. Note that this set, in the case at hand, is group specific, indicated by the jt index in $J_{0,jt}$. Chernozhukov and Hong (2002) propose choosing the trimming constant k such that a constant fraction of individuals is discarded, i.e.,

$$\frac{\#J_{0,jt}(k)}{\#J_{0,jt}(0)} = (1 - q_0) \times 100\%,$$

where $q_0 = 0.1$ has been found to work well in simulations (cf., Chernozhukov et al., 2015).

Step 2 (Chernozhukov and Hong, 2002). Using the set of observations in $J_{0,jt}(k)$, the second step entails running a standard quantile regression (Koenker and Bassett Jr., 1978) to obtain a first (inefficient) estimate of the censored quantile regression coefficients:

$$\widehat{\gamma}_{jt}^{0}(\tau) = \arg\min_{d \in \mathbb{R}^K} \sum_{i \in J_{0,jt}(k)} \rho_\tau\left(w_{ijt} - z_{ijt}^{\intercal} d\right). \tag{C.2}$$

Using these estimates, the following set is obtained:

$$J_{1,jt} = \{i : z_{ijt}^{\intercal} \widehat{\gamma}_{0,jt}(\tau) < c_{jt} - \kappa(\tau)\}, \tag{C.3}$$

[21] Specifically, to estimate the probability of not being censored, the specification includes a full set of educational attainment and age group dummies, as well as a male dummy. Additionally, it incorporates interactions between the male dummy and both the educational attainment and age group dummies. Fully interacting all three categorical variables with one another turned out to be too flexible given the sometimes comparatively small number of worker-level observations.

with $\rho_\tau(\cdot)$ being the check function, and $\kappa(\tau)$ being a trimming constant serving the same purpose as k in the first step. As noted in Chernozhukov et al. (2015, 2019), $\kappa(\tau)$ should also be chosen such that a constant fraction of $J_{1,jt}$ is discarded. In practice, this is achieved by setting $\kappa(\tau)$ to be the q_1th quantile of the set of individuals defined by $c_{jt} - z_{ijt}^\top \widehat{\gamma}_{jt}^0(\tau)$ conditional on $z_{ijt}^\top \widehat{\gamma}_{jt}^0(\tau) < c_{jt}$. A value which worked well in practice is $q_1 = 0.03$ (Chernozhukov et al., 2015, 2019) which is also used in the empirical analysis here. This step aims at bolstering the efficiency of the final third step.

Step 3 (Chernozhukov and Hong, 2002). Lastly, the obtained set $J_{1,jt}$ is used for estimating the parameters from the model specified in (4.1), i.e.,

$$\widehat{\gamma}_{jt}^1(\tau) = \arg\min_{d \in \mathbb{R}^K} \sum_{i \in J_{1,jt}} \rho_\tau\left(w_{ijt} - z_{ijt}^\top d\right). \tag{C.4}$$

Additional computational details

The parameters are estimated using the `cqiv` Stata command developed by Chernozhukov et al. (2019). To reduce runtime and because the full variance-covariance matrix is not needed in the first-step group-specific quantile regressions, the `qrprocess` Stata command proposed by Chernozhukov et al. (2022, 2020) is used. This allows for efficient computation of the parameter estimates while skipping the time-consuming calculation of the variance-covariance matrices.

C.2 Between- and within-inequality: Additional remarks

Since the empirical analysis largely follows the approach outlined by Oka and Yamada (2023), it is worthwhile to provide some comments about specific aspects of their procedure that differ from our approach.

The findings in Oka and Yamada (2023) imply that the policy-induced change of an *isolated* return to a given observable worker characteristic – and thus also its respective effect on between wage inequality – actually *depends on the particular group* that one considers. For example, their results imply that the effect of changes in the minimum wage affected the gender wage differential much more for prototypical individuals with 12 years education and 5 years experience than for prototypical individuals with 16 years of education and 10 years of experience (Oka and Yamada, 2023, Figure 12, p.355). The latter would be equivalent to claiming that the effect on the return to gender is different for group A than for group B.

The following shows that this is not the case as long as there are no interactions in the first-stage quantile regressions. Start by considering (4.14) for a special prototypical type-A and type-B worker that differ only in the kth element:

$$\bar{z}_{A,\cdot k}^{\top} \equiv \begin{pmatrix} \bar{z}_1 & \cdots & z_{A,k} & \cdots & \bar{z}_K \end{pmatrix}, \quad \bar{z}_{B,\cdot k}^{\top} \equiv \begin{pmatrix} \bar{z}_1 & \cdots & z_{B,k} & \cdots & \bar{z}_K \end{pmatrix}.$$

Inserting these two vectors in (4.14), it is immediately apparent that the difference that arises due to the group-level effect on the kth covariate is independent of the group one considers *if one holds everything but the kth element fixed* – something that has also been highlighted in Xu et al. (2023). Concretely, for these particular vectors of worker characteristic, the isolated effect of the CBA coverage drop – similarly to (4.14) – reads

$$\Delta^{f,cf}(\tau \,|\, \bar{z}_{B,\cdot k}) - \Delta^{f,cf}(\tau \,|\, \bar{z}_{A,\cdot k}) = \begin{pmatrix} \underset{(1\times K)}{\bar{z}_{B,\cdot k}^{\top}} - \underset{(1\times K)}{\bar{z}_{A,\cdot k}^{\top}} \end{pmatrix} \left[\left(\mathbb{E}[u_{jt_1}] - \mathbb{E}[\bar{u}_{jt_0}] \right) \mathbf{I}_K \, \beta(\tau) \right].$$

Considering these particular vectors of worker characteristics implies that the *isolated* policy-induced effect on the wage differential, due to differences in the kth covariate, is given by:

$$\begin{pmatrix} \underset{(1\times K)}{\bar{z}_{B,\cdot k}^{\top}} - \underset{(1\times K)}{\bar{z}_{A,\cdot k}^{\top}} \end{pmatrix} = \begin{pmatrix} \bar{z}_1 \\ \vdots \\ z_{B,k} \\ \vdots \\ \bar{z}_K \end{pmatrix} - \begin{pmatrix} \bar{z}_1 \\ \vdots \\ z_{A,k} \\ \vdots \\ \bar{z}_K \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ z_{B,k} - z_{A,k} \\ \vdots \\ 0 \end{pmatrix}.$$

The latter formally replicates the intuitive result that the effect of isolated returns to worker observables and the corresponding effect of CBA coverage are *not* group-dependent, i.e., they remain the same regardless of which other covariates are held fixed.

What drives the group dependency in Oka and Yamada (2023) in their analysis of minimum wage effects in the US? Upon closer examination, it turns out that they replace the quantile prediction with the binding minimum wage level whenever the predicted quantile for a prototypical worker type is less than the relevant minimum wage rate. This induces an additional source of variation, disproportionately affecting groups with lower predicted quantiles (e.g., less experienced/educated female workers), which in turn leads to group-dependent differences in the induced policy effect. If there is no need to additionally impute or replace values due to factors such as binding minimum wages, the wage differential arising from policy-induced shifts in returns to observables is, as shown above, not dependent on the group. A similar point has been made in Xu et al. (2023). Intuitively, this must be the case as long as the underlying model for the conditional quantiles does not include interactions between worker observables, as in our empirical analysis.

C.3 Estimation of national means of inequality measures

In the following, further details regarding the estimation of involved parameters are presented. All estimations make use of either the individual *LIAB* worker-level weights, henceforth ω_{ijt}, or the group-level aggregated version of them, i.e., $\psi_{jt} \equiv \sum_{i\in(j,t)} \omega_{ijt}$. For later reference, denote the normalized version of ψ_{jt} as $\widetilde{\psi}_{jt} \equiv \psi_{jt} / \left(\sum_{(j,t)} \psi_{jt}\right)$.

Joint variance-covariance matrix

The between-group and within-group inequality measures, formalized in (4.14) and (4.17) respectively, require obtaining valid standard errors for any linear combination of the involved returns from (4.2), i.e., $\beta(\tau)$. Specifically, this involves estimating the joint variance-covariance matrix across all relevant subsets of equations, with the exact subset determined by the worker types considered. This, in turn, allows for the computation of standard errors of any linear combinations that are implied by the between- or within-inequality measures. Since the two-step procedure by Chetverikov et al. (2016) allows for cluster sampling at the second stage, the variance-covariance matrix for the coefficient of the second-stage system of equations takes clustering at the level of sector-regions into account.

For simplicity, the following analysis considers the full system involving all K equations rather than just a subset. Any subset consideration is immediately implied by this. Define the stacked estimates for the entire system of second-stage coefficients in (4.3) as $\widehat{\boldsymbol{\Theta}}(\tau)$, i.e., $\widehat{\boldsymbol{\Theta}}(\tau)$ comprises both $\widehat{\boldsymbol{\theta}}(\tau)$ and $\widehat{\beta}(\tau)$. As to the involved dimensions, denote the pooled sample across all sector-regions, j, and over time, t, as M, and the number of observations within a sector-region j as M_j. The vector of stacked coefficients is of dimension $(K(Q+1)\times 1)$. The stacked matrix of second-stage covariates is referred to as $\widetilde{X}$ with dimension $(KM\times K(Q+1))$. Correspondingly, the $(KM_j\times K(Q+1))$ matrix $\widetilde{X}_j$ contains only the second-stage observations of the involved covariates within sector-region j. Lastly, the diagonal $(KM\times KM)$ dimensional matrix of aggregated sampling weights is denoted as Ψ. Accordingly, the diagonal $(M_j\times M_j)$ dimensional weighting matrix comprising only observations from sector-region j is denoted as Ψ_j. The asymptotic variance-covariance matrix, which takes into account cross-equation relationships and allows for testing hypotheses on linear combinations of coefficients from different equations in the system of K equations, is given by:

$$\underset{(K(Q+1)\times K(Q+1))}{\widehat{AVar}(\widehat{\boldsymbol{\Theta}}(\tau))} = \left(\widetilde{X}^{\top}\Psi\widetilde{X}\right)^{-1}\left(\sum_j\left(\widetilde{X}_j^{\top}\Psi_j^{1/2}\hat{e}_j\right)\left(\widetilde{X}_j^{\top}\Psi_j^{1/2}\hat{e}_j\right)^{\top}\right)\left(\widetilde{X}^{\top}\Psi\widetilde{X}\right)^{-1}, \quad \text{(C.5)}$$

with $\hat{e}_j$ being the $(KM_j\times 1)$ dimensional vector of residuals, i.e., $\hat{e}_j$ contains residuals for observations in the jth cluster.

In practice, this joint and stacked variance-covariance matrix is obtained by first estimating the parameters of all second-stage regressions as specified in (4.2) using weighted least squares with group-aggregated weights, ψ_{jt}. In the second step, the joint variance-covariance matrix of the desired linear combination involving a subset of the K second-stage equations is computed to obtain valid standard errors for the linear combination of interest. To achieve this, the cluster-adjusted procedure implemented in the `suest` Stata command (Weesie, 2000) is used to obtain the asymptotic variance-covariance matrix from (C.5).

Estimation of inequality measures

Since all the introduced inequality measures reduce to a combination of the term in (4.11) or the factual temporal differences specified in (4.12), it is only necessary to discuss the estimation of these two parameters.

Essentially, the term in (4.11) is a prediction of a parameter that depends on the deterministic vector of worker characteristic $\bar{z}$. Obtaining this prediction requires two things: (i) estimates for the K-dimensional vector $\beta(\tau)$ along with valid standard errors for estimating the joint variance-covariance matrix, and (ii) an estimate for the expectation of the group-specific CBA coverage rates. Note that the term in (4.11) is a linear combination of a subset of the K second-stage regression coefficients, $\beta(\tau)$. This subset is dictated by the deterministic worker-characteristic vector $\bar{z}$. Furthermore, the term that describes the CBA difference is treated as a quasi-deterministic scalar as well, with the difference in CBA coverage shares over time being computed using a weighted average, i.e.,

$$\widehat{\mathbb{E}}[u_{jt_1}] = \sum_{j \in t_1} u_{jt_1} \cdot \widetilde{\psi}_{jt_1}, \quad \widehat{\mathbb{E}}[u_{jt_0}] = \sum_{j \in t_0} u_{jt_0} \cdot \widetilde{\psi}_{jt_0},$$

$$\implies \overline{\Delta}^{t_1,t_0}(u_j) \equiv \left(\sum_{j \in t_1} u_{jt_1} \cdot \widetilde{\psi}_{jt_1} \right) - \left(\sum_{j \in t_0} u_{jt_0} \cdot \widetilde{\psi}_{jt_0} \right).$$

Given the very large sample size, the variability of $\overline{\Delta}^{t_1,t_0}(u_j)$ is ignored, i.e., $\overline{\Delta}^{t_1,t_0}(u_j)$ is effectively treated as a deterministic scaling parameter.

Thus, the point estimate for $\Delta^{f,cf}(\tau \mid \bar{z})$ in (4.11) and its standard error are obtained by considering suitable linear combinations of the involved point estimates. Following the procedure that yields (C.5), the standard error for this particular linear combination governed by $\bar{z}$ can be estimated. Ultimately, the point estimate arises as

$$\widehat{\Delta}^{f,cf}(\tau \mid \bar{z}) = \bar{z}^{\mathsf{T}} \left(\left[\left(\overline{\Delta}^{t_1,t_0}(u_j) \right) \mathbf{I}_K \right] \widehat{\beta}(\tau) \right).$$

As demonstrated in the main text, the isolated impact of the CBA coverage drop on both the between- and within-inequality measures are just suitably combined versions of (4.11), i.e., the point estimates of these measures as well as the standard error computation follow immediately from this discussion. In practice, standard errors for the between-inequality measure, as in (4.14), and for the within-inequality measure, as in (4.17), are computed using the `lincom` command in Stata.

Estimation of factual temporal differences

The last required element concerns the factual evolution of the considered parameters. Note that the functional form of the first-stage (factual) quantiles is given in (4.1). Using suitable first-stage parameter estimates, the factual temporal difference from period t_0 to period t_1 is given by

$$\Delta^{f}_{t_1,t_0}(\tau\,|\,\bar{z}) = \bar{z}^{\intercal}\left(\mathbb{E}[\widehat{\gamma}_{jt_1}(\tau)] - \mathbb{E}[\widehat{\gamma}_{jt_0}(\tau)]\right).$$

Being a discrete difference over time, this can be expressed in terms of a saturated linear regression:

$$\begin{aligned}\mathbb{E}[\widehat{\gamma}_{jt}(\tau)\,|\,D_t] &= \alpha(\tau) + \sum_{t=t_0+1}^{t_1} \phi_t(\tau)\, D_t,\\ \implies \mathbb{E}[\widehat{\gamma}_{jt_1}(\tau)] - \mathbb{E}[\widehat{\gamma}_{jt_0}(\tau)] &= \phi_{t_1}(\tau)\end{aligned}$$

with D_t being a year dummy equal to 1 in t and 0 otherwise. Hence, a point estimate for the change over time from the base year, t_0, to t_1 can be obtained using the estimate of $\phi_{t_1}(\tau)$. Lastly, note that all this refers to the stacked considerations involving all K first-stage coefficients. Estimation is carried out similarly to the above, using weighted least squares and employing the cluster-adjusted procedure described in Weesie (2000).

For estimating the joint variance-covariance matrix of a discrete combination of the K coefficients governed by $\bar{z}$, the same procedure as outlined above for $\widehat{\Delta}^{f,cf}(\tau\,|\,\bar{z})$ can be employed. All parameters that concern the temporal evolution of either the within- or between-inequality measures are just suitable combinations of the latter, i.e.,

$$\begin{aligned}\Delta^{B,f}_{t_1,t_0}(\tau\,|\,\bar{z}_A,\bar{z}_B) &= \left(\bar{z}_B^{\intercal} - \bar{z}_A^{\intercal}\right) \underset{(K\times 1)}{\phi_{t_1}(\tau)},\\ \Delta^{W,f}_{t_1,t_0}(\tau_A,\tau_B\,|\,\bar{z}) &= \bar{z}^{\intercal}\left(\underset{(K\times 1)}{\phi_{t_1}(\tau_B)} - \underset{(K\times 1)}{\phi_{t_1}(\tau_A)}\right).\end{aligned}$$

C.4 Recovering the unconditional distributional effect

In the following, the procedure according to which draws from the unconditional distribution of earnings are simulated is described. Beyond, figure C.1 displays the procedure's successive steps graphically. The procedure is based on the insights formulated in Machado and Mata (2005) and is essentially an application of the probability integral transformation. As intuitively described in Machado and Mata (2005) and Autor et al. (2005), the following steps – adjusted for the current case – are involved in sampling from the unconditional distribution of earnings:

1. Estimate the parameters of the conditional quantile regression model as defined in (4.1). Parameter estimates are obtained by means of (4.6), i.e., estimates incorporate the 3-step censored quantile regression model algorithm proposed by Chernozhukov and Hong (2002).

   1.1 To take into account the variability of the CBA coverage effect from the second stage, an exchangeable bootstrap procedure is applied (Chernozhukov et al., 2013; Van Kerm, 2022). In particular, sampling weights for a given cluster are drawn from a standard exponential distribution and are assigned to observations from the sample of first-step regression coefficients (for each τ). In total, this procedure is carried out B times.

   1.2 Use the bootstrap sample weights of the bth draw of the panel of first-stage regression coefficients to estimate the CBA coverage effect as in (4.4). Denote the estimate of the CBA coverage effect in the third episode that stems from the bth bootstrap sample as $\widetilde{\widehat{\beta}}_b(\tau) = \widehat{\beta}_{b,0}(\tau) + \widehat{\beta}_{b,2}(\tau)$, with $b = 1, \ldots, B$.

   1.3 Using $\widetilde{\widehat{\beta}}_b(\tau)$, generate the counterfactual return for a given group, $\widehat{\gamma}^{cf,b}_{j,2014}(\tau)$ as in (4.8) for each τ.

2. For each $b = 1, \ldots, B$, carry out the following steps:

   2.1 Randomly draw $s_b = 1, \ldots, S_b$ observations (with replacement) from the entire worker-level sample to obtain a random sample of worker characteristics. The b index highlights that this sampling is carried out separately for each bootstrap repetition $b = 1, \ldots, B$ from step 1. A given sampled observation s_b is denoted as $z^{s_b}_{ijt}$.

   2.2 For each sampled worker-level observation, generate a random draw $\tau^{s_b}_i \sim U(0,1)$, and further adjust these draws to match the closest value within the predefined set $\mathscr{S} = \{0.05, 0.10, \ldots, 0.90, 0.95\}$, to manage computational efficiency given the extensive number of overall group-level quantile regressions.

2.3 Given the draws of $z_{ijt}^{s_b}$ and $\tau_i^{s_b}$, sample realizations are simulated as

$$\{w_{ijt}^{b} = \left(z_{ijt}^{s_b}\right)^{\intercal} \widehat{\gamma}_{jt}(\tau_i^{s_b})\}_{s_b=1}^{S_b},$$

with w_{ijt}^{b} denoting the simulated realization that is associated with exchangeable bootstrap weights of the bth draw. The sampling procedure is designed such that the number of samples drawn is dependent on both the group and time index. In this case, the number of draws is chosen so that the average number of draws per sector-region equals 500 if the sector-regions in the *LIAB* data were equally sized. Since some sectors are larger than others in a given year, this approach ensures that the draws at the worker level accurately reflect the proportional sizes of the groups for each year in the *LIAB* sample.

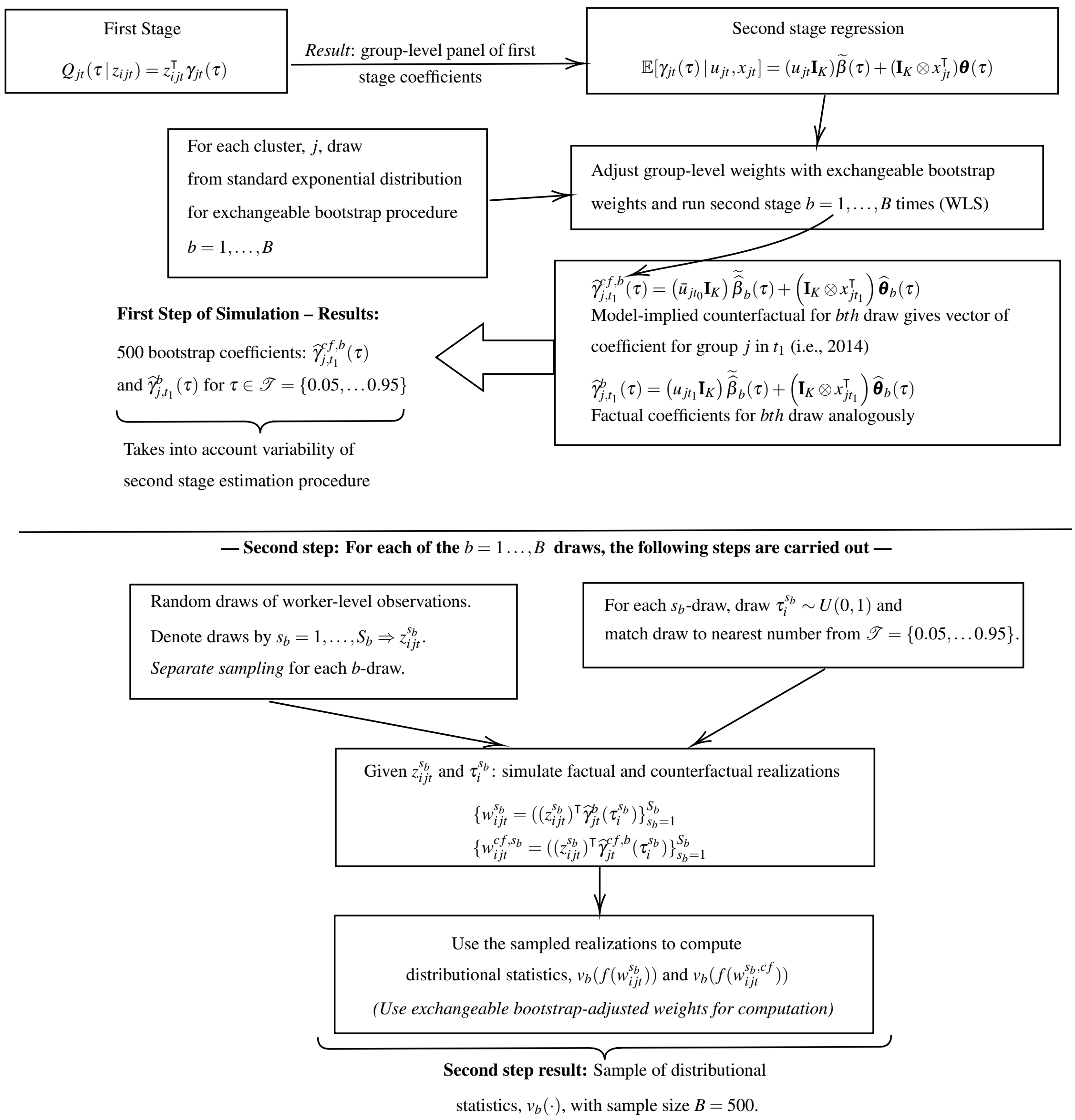
2.4 The unconditional counterfactual distribution in 2014 that incorporates the thought experiment of holding the CBA coverage level fixed at its 1996 level is simulated by using the counterfactual coefficients as given in (4.8), i.e.,

$$\{w_{ij,2014}^{cf,s_b} = \left(z_{ij,2014}^{s_b}\right)^{\intercal} \widehat{\gamma}_{j,2014}^{cf,b}(\tau_i^{s_b})\}_{s_b=1}^{S_b},$$

with $\widehat{\gamma}_{j,2014}^{cf,b}(\tau_i^{s_b})$ being the estimate using the bth bootstrap draw. Analogously to the case above, denote the simulated counterfactual realization that is associated with the bth draw's exchangeable bootstrap weights as $w_{ijt}^{cf,b}$.

3. Using the sampled realizations of both the factual and counterfactual processes, the distributional statistics $v_b(f(w_{ijt}^{b}))$ and $v_b(f(w_{ijt}^{cf,b}))$ can be obtained. For the computation of these parameters, sampling weights adjusted with the b-specific exchangeable bootstrap weights are used to account for the bootstrap sampling described in the first-stage estimation step.

Figure C.1: Simulation procedure for unconditional effects



Notes: Description of the simulation procedure that is used to simulate draws from the unconditional distribution of earnings and to compute distributional statistics from this simulated sample as described in detail in appendix C.4.

As pointed out by Autor et al. (2005), this sampling procedure is equivalent to integrating over the distribution of τ and z and ultimately results in a sample that is – in the case at hand – either drawn from the factual or counterfactual unconditional distribution of log daily earnings:

$$f(w_{ijt}) = \int_\tau \int_z \widehat{Q}_{ijt}(\tau \,|\, z_{ijt})\; g(z)\; dz d\tau,$$
$$f(w^{cf}_{ij,2014}) = \int_\tau \int_z \widehat{Q}^{cf}_{ij,2014}(\tau \,|\, z_{ij,2014})\; g(z)\; dz d\tau,$$

with $g(z)$ being the distribution of covariates.

Intuitively, the CBA coverage effect is computed $B = 500$ times. For *each of these draws*, a sample of unconditional log earnings is simulated by means of the procedure described above. In the process of doing so, the sampling of worker-level observations *per sector-region in a given year* amounts to, on average, $S_b = 500$.[22] This is equivalent to running the Machado and Mata (2005) sampling procedure B times. Each time, a new random sample is drawn, such that, ultimately, this procedure takes into account both the variability of the CBA coverage effect from the second stage regression from (4.4) as well as the sampling procedure's variability.

Using the bootstrap resamples, confidence bounds in the spirit of a percentile bootstrap approach (see, for example Firpo and Pinto, 2016) can be constructed. Define the αth percentile of the ordered sample of $v_b(\cdot)$ with $b = 1, \ldots, B$ as $v_{\langle \alpha \cdot B \rangle}$ such that, as B gets large:

$$\frac{1}{B} \sum_{b=1}^{B} \mathbb{1} \left[v_b \leq v_{\langle \alpha \cdot B \rangle} \right] \rightarrow \alpha, \text{ with } \alpha \in (0,1),$$

and $v_b(\cdot)$ referring to the distributional statistic computed with the bth repetition. Using

[22]The exact procedure boils down to drawing $(19 \times 11) \times 500 = 104{,}500$ worker level observations per year from the sample of worker level in a given year, with 19 being the number of sectors and 11 being the number of regions under consideration. This ensures that the proportion of sector-regions are upheld, with larger sector-regions being drawn more often than smaller sector-regions. In turn, the average number of worker-level observations in a given sector-region is 500, or, put differently, if the sector-regions were all equally sized, 500 worker-level observations would have been drawn per sector-region.

this, the percentile bootstrap confidence interval reads as

$$\mathrm{CI}\left(v(\cdot),(1-\alpha)\cdot 100\%\right)=\left[v_{\langle(\alpha/2)\cdot B\rangle},v_{\langle(1-\alpha/2)\cdot B\rangle}\right], \tag{C.6}$$

with CI being the approximate $(1-\alpha)\cdot 100\%$ percentile bootstrap confidence interval for the distributional statistic $v(\cdot)$. Given this nested bootstrap and resampling scheme, it is not possible to compute the distributional statistic using the 'actual sample' because everything is based on repeatedly simulated draws. Therefore, the reported effects for the distributional statistic $v(\cdot)$ correspond to the mean over all $B=500$ replicates, i.e,

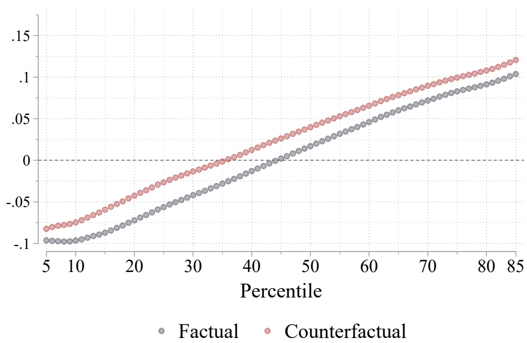
$$\overline{v}\equiv\frac{1}{B}\sum_{b=1}^{B}v_b(f(w_{ijt}^{b})), \tag{C.7}$$

$$\overline{v}^{cf}\equiv\frac{1}{B}\sum_{b=1}^{B}v_b(f(w_{ij,2014}^{cf,b})). \tag{C.8}$$

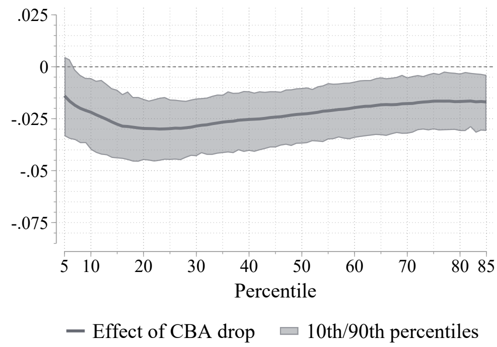
C.5 Additional figures

Figure C.2: Effect on unconditional distribution – temporal evolution and effect of CBA coverage drop (male and female sub-sample)

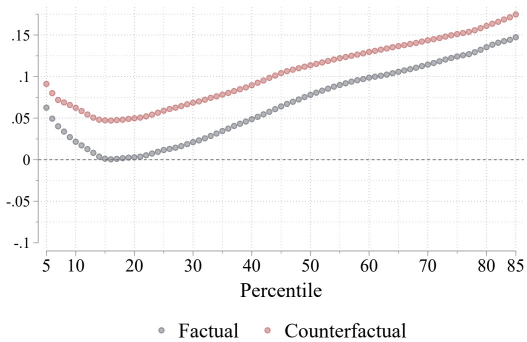
(a) Temporal change (male)



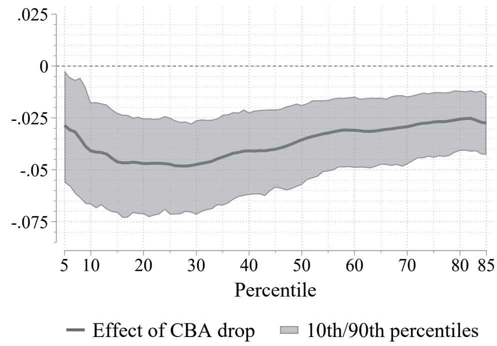
(b) CBA effect (male)



(c) Temporal change (female)

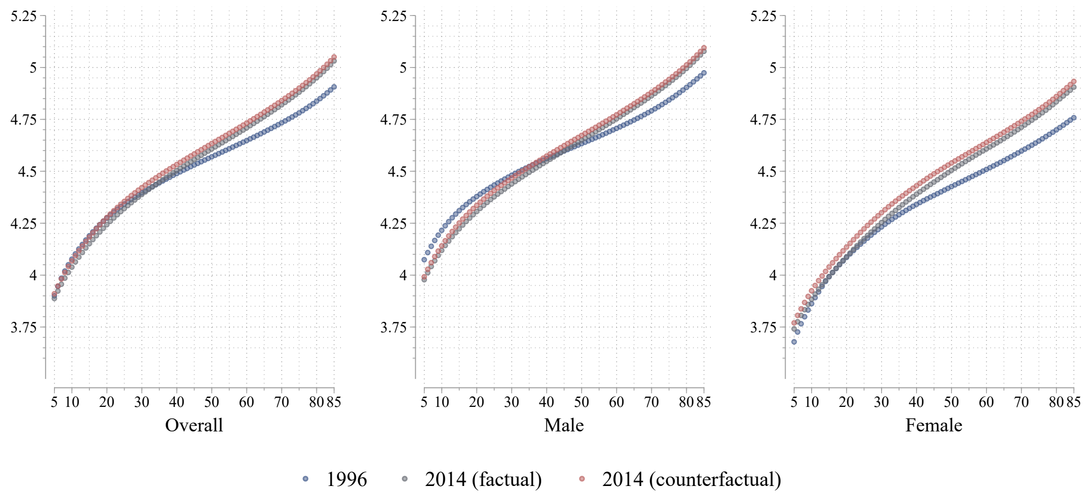


(d) CBA effect (female)



Notes: Subfigures C.2a and C.2c show the log point change in factual and counterfactual earnings from 1996 to 2014 for the the male and female sub-sample as given in equations (4.20) and (4.21) for several percentiles of the unconditional distribution of daily earnings. Subfigures C.2b and C.2d display the part of the respective difference in daily earnings that can be ascribed to the drop in CBA coverage from 1996 to 2014, as given in (4.22). The values are computed using data that are the result of the simulation-based approach, which is described in detail in appendix C.4. The isolated effects of the CBA coverage drop, indicated by the solid lines in subfigures C.2b and C.2d, correspond to the respective average over all simulated and bootstrapped worker-level observations for a given percentile value that are computed analogously to (C.7) and (C.8). The shaded area in subfigures C.2b and C.2d correspond to the 10th and 90th percentiles of the sample of computed bootstrap values of the effect. These values thus represent the bounds of the 90% percentile bootstrap confidence intervals, as in (C.6) and as described, e.g., in Firpo and Pinto (2016).
Source: Simulated data based on *LIAB QM2 9319*, own calculations.

Figure C.3: Quantile functions: 1996 vs. 2014



Notes: Figure shows the quantile functions for the overall sample, as well as for the male and female subsample of the unconditional distribution of log daily earnings. The values are computed using data that are the result of the simulation-based approach, which is described in detail in appendix C.4. The displayed values refer to averages as in (C.8).
Source: Simulated data based on *LIAB QM2 9319*, own calculations.

C.6 Additional tables

Table C.1: Group means of returns to first-stage worker characteristics for various conditional quantiles

	1996	2002	2008	2014
Returns (group-mean)				
Constant				
10	4.285	4.276	4.170	4.228
15	4.358	4.357	4.257	4.306
20	4.409	4.417	4.325	4.369
25	4.453	4.468	4.383	4.420
30	4.487	4.512	4.428	4.466
35	4.524	4.548	4.473	4.507
40	4.555	4.587	4.512	4.547

Continued on next page

Table C.1 – continued from previous page

	1996	2002	2008	2014
45	4.588	4.621	4.552	4.588
50	4.623	4.655	4.593	4.624
55	4.652	4.690	4.634	4.664
60	4.686	4.722	4.672	4.704
65	4.722	4.756	4.713	4.744
70	4.762	4.793	4.758	4.787
75	4.805	4.832	4.801	4.836
80	4.848	4.879	4.849	4.882
85	4.905	4.931	4.917	4.946
90	4.962	4.996	5.009	5.017
Female				
10	-0.336	-0.410	-0.396	-0.257
15	-0.296	-0.341	-0.338	-0.225
20	-0.268	-0.298	-0.295	-0.212
25	-0.237	-0.268	-0.260	-0.192
30	-0.221	-0.245	-0.231	-0.176
35	-0.208	-0.226	-0.214	-0.163
40	-0.197	-0.205	-0.194	-0.150
45	-0.182	-0.189	-0.179	-0.138
50	-0.174	-0.177	-0.166	-0.128
55	-0.170	-0.166	-0.153	-0.125
60	-0.168	-0.158	-0.142	-0.116
65	-0.166	-0.150	-0.130	-0.107
70	-0.161	-0.144	-0.121	-0.096
75	-0.158	-0.136	-0.113	-0.092
80	-0.154	-0.134	-0.109	-0.083
85	-0.149	-0.128	-0.108	-0.081
90	-0.139	-0.120	-0.096	-0.076
No formal education				
10	-0.180	-0.274	-0.245	-0.266
15	-0.176	-0.257	-0.207	-0.222
20	-0.171	-0.227	-0.204	-0.209
25	-0.154	-0.216	-0.183	-0.204
30	-0.148	-0.207	-0.176	-0.196

Continued on next page

Table C.1 – continued from previous page

	1996	2002	2008	2014
35	-0.142	-0.197	-0.170	-0.190
40	-0.138	-0.194	-0.167	-0.187
45	-0.145	-0.194	-0.169	-0.186
50	-0.147	-0.190	-0.171	-0.186
55	-0.148	-0.188	-0.166	-0.185
60	-0.149	-0.184	-0.168	-0.184
65	-0.158	-0.183	-0.170	-0.188
70	-0.163	-0.185	-0.172	-0.188
75	-0.166	-0.183	-0.178	-0.187
80	-0.168	-0.184	-0.176	-0.190
85	-0.166	-0.181	-0.173	-0.172
90	-0.157	-0.175	-0.168	-0.161
University education				
10	0.360	0.335	0.380	0.308
15	0.381	0.366	0.394	0.346
20	0.387	0.363	0.406	0.369
25	0.388	0.373	0.432	0.377
30	0.382	0.375	0.427	0.388
35	0.375	0.387	0.440	0.409
40	0.369	0.387	0.445	0.413
45	0.360	0.393	0.433	0.409
50	0.352	0.396	0.438	0.419
55	0.357	0.387	0.438	0.418
60	0.369	0.377	0.429	0.423
65	0.359	0.369	0.418	0.422
70	0.343	0.362	0.402	0.406
75	0.351	0.350	0.381	0.394
80	0.315	0.348	0.374	0.368
85	0.292	0.330	0.347	0.343
90	0.265	0.284	0.286	0.305
Age (18-25)				
10	-0.161	-0.212	-0.220	-0.179
15	-0.178	-0.211	-0.226	-0.190
20	-0.193	-0.216	-0.233	-0.198

Continued on next page

Table C.1 – continued from previous page				
	1996	2002	2008	2014
25	-0.207	-0.225	-0.245	-0.199
30	-0.216	-0.232	-0.249	-0.205
35	-0.228	-0.240	-0.255	-0.211
40	-0.231	-0.252	-0.263	-0.219
45	-0.239	-0.258	-0.270	-0.228
50	-0.248	-0.269	-0.280	-0.235
55	-0.244	-0.275	-0.288	-0.246
60	-0.251	-0.281	-0.300	-0.257
65	-0.259	-0.288	-0.310	-0.265
70	-0.272	-0.294	-0.324	-0.275
75	-0.282	-0.304	-0.332	-0.288
80	-0.287	-0.317	-0.342	-0.297
85	-0.306	-0.327	-0.360	-0.319
90	-0.319	-0.341	-0.382	-0.335
Age (26-35)				
10	-0.039	-0.041	-0.065	-0.068
15	-0.047	-0.047	-0.078	-0.074
20	-0.053	-0.054	-0.087	-0.083
25	-0.062	-0.063	-0.097	-0.090
30	-0.064	-0.063	-0.100	-0.094
35	-0.068	-0.068	-0.106	-0.100
40	-0.067	-0.073	-0.112	-0.106
45	-0.070	-0.077	-0.112	-0.111
50	-0.074	-0.081	-0.119	-0.116
55	-0.077	-0.084	-0.123	-0.121
60	-0.081	-0.087	-0.127	-0.128
65	-0.087	-0.089	-0.128	-0.132
70	-0.093	-0.089	-0.137	-0.136
75	-0.098	-0.092	-0.135	-0.143
80	-0.098	-0.100	-0.140	-0.144
85	-0.107	-0.103	-0.148	-0.149
90	-0.105	-0.109	-0.148	-0.153
Age (51-64)				
10	0.015	-0.003	0.007	-0.005
15	0.021	0.010	0.013	0.005

Continued on next page

Table C.1 – continued from previous page				
	1996	2002	2008	2014
20	0.024	0.015	0.014	0.002
25	0.025	0.016	0.015	0.009
30	0.027	0.022	0.015	0.009
35	0.024	0.023	0.016	0.010
40	0.027	0.017	0.011	0.011
45	0.024	0.018	0.017	0.011
50	0.024	0.017	0.016	0.011
55	0.031	0.018	0.014	0.012
60	0.034	0.020	0.015	0.010
65	0.039	0.021	0.020	0.011
70	0.043	0.026	0.018	0.011
75	0.046	0.028	0.021	0.013
80	0.044	0.035	0.026	0.022
85	0.042	0.032	0.024	0.028
90	0.031	0.034	0.023	0.026

Notes: The table shows the weighted group-mean of returns to the first-stage worker characteristics across several conditional quantiles. Results are presented for four selected years, representing the boundaries defining the episodes considered in (4.4). The weighted group-mean is calculated by computing the mean of returns for the groups considered in the second-stage panel data regression, using aggregated worker-level *LIAB* weights as the weighting factors. All coefficients are relative to the base group (male, vocational training, aged 36-50).
Data: LIAB QM2 9319, own calculations.

Chapter 5

Dissertation Summary and Conclusion

From the mid-1990s to the mid-2010s, the post-reunification German economy underwent some major structural changes. A key feature of this development was an unprecedented decentralization of wage-setting mechanisms in the form of de-unionization. This coincided with a stark increase in wage and earnings inequality and, moreover, implied the elimination of collectively bargained wage floors for many workers. Against the backdrop of this deterioration of wage floors, a federally binding minimum wage was introduced in 2015, which was undoubtedly a massive labor market intervention because of the scale with which it affected the workforce. Chapter 2 and chapter 4 of this dissertation empirically examined the distributional effects of these two important labor market institutions in Germany and demonstrated the application of innovative econometric techniques. Furthermore, chapter 3 critically assessed the potential of an econometric approach for estimating quantile treatment effects that combines recentered influence function regressions with a difference-in-differences idea, which has been suggested for estimating the distributional effect of, e.g., the minimum wage introduction in Germany in previous contributions.

In the study presented in chapter 2, we paint a comprehensive picture of the causal distributional effect of the minimum wage. Unlike other contributions that used small-scale survey data to consider the policy's effect on hourly wages (e.g., Caliendo et al., 2023), or were restricted to monthly earnings due to limitations in administrative data sources (Bossler and Schank, 2023), we combine two large-scale German databases to overcome these limitations. This, in turn, allows us to consider the impact of the minimum wage on the distributions of hourly earnings, monthly earnings, and hours worked. Crucially, by considering all three dimensions and making use of reliable, large-scale data, our results can help reconcile the partly conflicting findings regarding the distributional effect of the minimum wage introduction in Germany from previous studies (Bossler and Schank, 2023; Caliendo et al., 2023). Adding to the existing literature, our findings suggest that the introduction of the minimum wage causally led to substantial distributional effects along the dimensions of hourly wages and monthly earnings, whereas only small effects on hours worked were found for the group of low-wage workers. Building on these findings, our study further demonstrates that the introduction of the minimum wage can explain much of the recent decrease in inequality in both hourly pay and monthly earnings. Another important aspect of our study is the methodological approach we employ. Concretely, we combine a difference-in-differences with a distribution regression approach

(Chernozhukov et al., 2013) to construct a counterfactual distribution absent treatment. In that way, we avoid the shortcomings of another approach that is built on recentered influence functions, around which the study in chapter 3 revolved. Our study further underscores the importance of taking account of the dynamic evolution of wage and earnings distributions in the years leading up to the introduction of the minimum wage. Acknowledging these pre-trends in our analysis, we demonstrate that effects would have been overstated if we did not correct for these systematic pre-trends. In conclusion, our findings have important implications for the ex post assessment of the newly introduced German minimum wage, contribute to painting a fuller picture of the policy's effect, and help to resolve previously conflicting findings regarding the policy's distributional effects.

While the proposed approach in chapter 2 centered around distribution regression techniques, several other studies have employed an alternative method that combines recentered influence function regressions (Firpo et al., 2009) with a difference-in-differences approach to estimate the distributional effects of the minimum wage introduction in Germany (Bossler and Schank, 2023; Caliendo et al., 2023; Bossler et al., 2024). Against the background of the increasing popularity of this approach, chapter 3 delved into its inherent shortcomings. Specifically, I show both formally and within a Monte Carlo simulation study that estimates obtained using a 'pooled RIF-DiD' approach for quantiles do not coincide with an estimate of the quantile treatment effect on the treated, which is the parameter of interest. Rather, the resulting estimate corresponds to a spurious effect approximated around a misleading quantile. Building on earlier contributions (Havnes and Mogstad, 2015; Rios-Avila and Maroto, 2024), I further propose two approaches to overcome these shortcomings. These approaches adjust the RIF-based difference-in-differences procedure so that treatment effects are approximated around the quantile of the relevant sub-population. Moreover, to connect the estimation strategies of the two suggested alternative RIF-DiD approaches with theoretical identification results, I clearly link these strategies to the required identifying assumptions formulated in previous contributions (Athey and Imbens, 2002, 2006; Roth and Sant'Anna, 2023). I further corroborate the formal and simulation findings by revisiting a real-world application (Hoynes and Patel, 2018), where I find that the differences between the misleading and the two suggested RIF-DiD approaches are often substantial. Concluding, the results of the study in chapter 3 underscore the importance of a thorough discussion on the appropriateness of the chosen econometric technique for estimating an identified object of interest, and might offer guidance regarding methodological choices in future studies.

The analysis in chapter 2 focused on evaluating a newly introduced labor market institution, the minimum wage. Conversely, chapter 4 presented a study that examines the heterogeneous distributional effects of a long-standing labor market institution in Germany which is considered a cornerstone of the German economy, i.e., union-mediated wage setting in the form of collective bargaining. In chapter 4, I broaden the large literature on the distributional effects of union-mediated wage setting in Germany by several aspects. On the one hand, I employ a novel methodological approach that takes the institutional setting in Germany into account. In particular, to align the econometric approach with the institutional setting, I utilize a grouped quantile regression approach (Chetverikov et al., 2016; Oka and Yamada, 2023) that exploits variation at the sector-region level – the level at which most collective agreements in Germany apply. When exploring the distributional consequence of de-unionization, previous studies conceived de-unionization as an aggregate compositional effect (e.g., Firpo et al., 2018; Biewen and Seckler, 2019; Baumgarten et al., 2020). In contrast, I consider a more nuanced effect and examine the extent to which shifts in collective bargaining agreement coverage altered the earnings structure at various relative positions of the worker productivity distribution. In doing so, I demonstrate that de-unionization between 1996 and 2014 substantially affected earnings differentials between and within worker types. Moreover, the findings presented in chapter 4 broaden the literature by proposing a novel way to estimate the effect of de-unionization on the unconditional distribution of earnings. Specifically, to recover the unconditional effect, I propose an innovative combination of the grouped quantile regression with a simulation-based approach proposed in Machado and Mata (2005), which, intuitively, is equivalent to modeling the interaction between the compositional and wage structure effects. Hence, I offer an additional lens through which the distributional impact of de-unionization in Germany can be assessed. In conclusion, my findings offer additional evidence on the significance of union-mediated wage setting and how de-unionization affected the German structure of earnings. Moreover, the findings presented in chapter 4 suggest that collective agreements constitute a critical institutional layer within the German labor market that mediates the remuneration for certain worker characteristics.

In summary, this dissertation is broadly characterized by two main themes. First, chapters 2 and 4 highlighted the importance of labor market institutions in shaping the distributions of wages and earnings. As a matter of fact, the two labor market institutions that were studied in this dissertation remain recurring themes in public discourse, underlining the persistent role they play in discussions outside academia as well. As to the minimum wage, the passionate debate about its increase in Germany at the time of writing this dissertation serves as an example (e.g., Göbel, 2024; Rossbach, 2024). On the other hand, against the background of employers withdrawing from the system of collective bargaining, the discussion about what measures could encourage employers to participate in collective negotiations continues unabated. For example, to create incentives for employers to apply collective bargaining agreements, one proposal by the ruling German federal government is to exclusively award public contracts to establishments bound by such agreements (Bundesministerium für Wirtschaft und Klimaschutz, 2023). Another example relates to the proposal by Barišić et al. (2023), who suggest that establishments bound by collective agreements should be granted simplified procedures for recruiting migrants from third countries. Hence, based on the topicality of the current discussion surrounding these two labor market institutions, I conclude that a thorough, empirically-based understanding of the impacts of these two important pillars of the German labor market is essential, to which this dissertation contributes.

The second general conclusion I draw from this dissertation relates to the second main theme addressed in chapters 2, 3, and 4, which refers to the importance of employing suitable econometric methods in empirical labor market research. As to the methodological choice in chapter 2, our decision to employ a distribution regression difference-in-differences approach was, among other things, guided by the fact that the minimum wage introduction implied mass points, and we required a methodological approach that could accommodate this. The econometric technique in chapter 4 revolved around considerations of how the specific German system of industrial relations can be accounted for when estimating the distributional effects of de-unionization in Germany. Lastly, chapter 3 most directly addressed the necessity of a careful examination of the employed econometric technique by highlighting the pitfalls involved in applying a RIF-DiD approach to estimate distributional treatment effects. In conclusion, this dissertation demonstrated the importance of thoroughly discussing the appropriateness of the chosen econometric techniques for estimating distributional effects that suit the requirements dictated, e.g., by specific institutional or data-driven features.

Bibliography

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies 72*(1), 1–19.

Acemoglu, D. and D. Autor (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, Volume 4, pp. 1043–1171. Elsevier.

Acemoglu, D. and P. Restrepo (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica 90*(5), 1973–2016.

Ahlfeldt, G. M., D. Roth, and T. Seidel (2018). The Regional Effects of Germany's National Minimum Wage. *Economics Letters 172*, 127–130.

Almond, D., H. W. Hoynes, and D. W. Schanzenbach (2011). Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes. *The Review of Economics and Statistics 93*(2), 387–403.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics*. Princeton University Press.

Antonczyk, D., T. DeLeire, and B. Fitzenberger (2018). Polarization and Rising Wage Inequality: Comparing the U.S. and Germany. *Econometrics 6*(2), 1–33.

Antonczyk, D., B. Fitzenberger, and K. Sommerfeld (2010). Rising Wage Inequality, the Decline of Collective Bargaining, and the Gender Wage Gap. *Labour Economics 17*(5), 835–847.

Athey, S. and G. W. Imbens (2002). Identification and Inference in Nonlinear Difference-in-Differences Models. *NBER Technical Working Paper t0280*.

Athey, S. and G. W. Imbens (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica 74*(2), 431–497.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–68.

Autor, D. H. and L. F. Katz (1999). Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*, Volume 3, pp. 1463–1555. Elsevier.

Autor, D. H., L. F. Katz, and M. S. Kearney (2005). Rising Wage Inequality: The Role of Composition and Prices. *NBER Working Paper Series 11628*.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics 90*(2), 300–323.

Autor, D. H., A. Manning, and C. L. Smith (2016). The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment. *American Economic Journal: Applied Economics 8*(1), 58–99.

Bachmann, R., H. Bonin, B. Boockmann, G. Demir, R. Felder, I. Isphording, R. Kalweit, N. Laub, C. Vonnahme, and C. Zimpelmann (2020). Auswirkungen des gesetzlichen Mindestlohns auf Löhne und Arbeitszeiten: Studie im Auftrag der Mindestlohnkommission. RWI Projektberichte, Essen. Additional information: Projektvergabe durch die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) - Vergabe-Nr. 536043 - Endbericht.

Barišić, M., S. Jäger, A. Manning, M. Muñoz, U. Rinne, and J. Stuhler (2023). Einwanderung klug, einfach und fair gestalten: Ein Vorschlag mit doppelter Dividende. *IZA Standpunkte Nr. 105*.

Bassier, I. (2021). The Impact of Centralized Bargaining on Spillovers and the Wage Structure in Monopsonistic Labour Markets. *WIDER Working Paper 2021/132*.

Baumgarten, D., G. Felbermayr, and S. Lehwald (2020). Dissecting Between-Plant and Within-Plant Wage Dispersion: Evidence from Germany. *Industrial Relations: A Journal of Economy and Society 59*(1), 85–122.

Bechmann, S., N. Tschersich, P. Ellguth, S. Kohaut, E. Baier, and C. Florian (2021). Technical Report on the IAB Establishment Panel – Wave 27 (2019). *FDZ-Methodenreport 1*(2021).

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Differences-In-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Bhaskar, V., A. Manning, and T. To (2002). Oligopsony and Monopsonistic Competition in Labor Markets. *Journal of Economic Perspectives 16*(2), 155–174.

Bhuller, M., K. O. Moene, M. Mogstad, and O. L. Vestad (2022). Facts and Fantasies About Wage Setting and Collective Bargaining. *Journal of Economic Perspectives 36*(4), 29–52.

Biasi, B. and H. Sarsons (2022). Flexible Wages, Bargaining, and the Gender Gap. *Quarterly Journal of Economics 137*(1), 215–266.

Biewen, M., B. Fitzenberger, and J. de Lazzer (2018). The Role of Employment Interruptions and Part-Time Work for the Rise in Wage Inequality. *IZA Journal of Labor Economics 7*(1), 1–34.

Biewen, M., B. Fitzenberger, and M. Rümmele (2022). Using Distribution Regression Difference-in-Differences to Evaluate the Effects of a Minimum Wage Introduction on the Distribution of Hourly Wages and Hours Worked. *IZA Discussion Paper No. 15534*.

Biewen, M. and A. Juhasz (2012). Understanding Rising Income Inequality in Germany, 1999/2000–2005/2006. *Review of Income and Wealth 58*(4), 622–647.

Biewen, M. and M. Seckler (2019). Unions, Internationalization, Tasks, Firms, and Worker Characteristics: A Detailed Decomposition Analysis of Rising Wage Inequality in Germany. *The Journal of Economic Inequality 17*(4), 461–498.

Blömer, M. J., M. Dolls, C. Fuest, M. Löffler, and A. Peichl (2015). German Public Finances through the Financial Crisis. *Fiscal Studies 36*(4), 453–474.

Bonhomme, S. and U. Sauder (2011). Recovering Distributions in Difference-in-Differences Models: A Comparison of Selective and Comprehensive Schooling. *The Review of Economics and Statistics 93*(2), 479–494.

Borgen, N. T., A. Haupt, and Ø. N. Wiborg (2022). Quantile Regression Estimands and Models: Revisiting the Motherhood Wage Penalty Debate. *European Sociological Review 39*(2), 317–331.

Bossler, M. (2017). Employment Expectations and Uncertainties Ahead of the New German Minimum Wage. *Scottish Journal of Political Economy 64*(4), 327–348.

Bossler, M. (2019). The Rise in Orientation at Collective Bargaining without a Formal Contract. *Industrial Relations: A Journal of Economy and Society 58*(1), 17–45.

Bossler, M. and H.-D. Gerner (2020). Employment Effects of the New German Minimum Wage: Evidence From Establishment-Level Microdata. *ILR Review 73*(5), 1070–1094.

Bossler, M., Y. Liang, and T. Schank (2024). The Devil Is in the Details: Heterogeneous Effects of the German Minimum Wage on Working Hours and Minijobs. *IZA Discussion Paper No. 16964*.

Bossler, M. and T. Schank (2023). Wage Inequality in Germany After the Minimum Wage Introduction. *Journal of Labor Economics 41*(3), 813–857.

Bound, J. and G. Johnson (1992). Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Hypotheses. *American Economic Review 82*(3), 371–392.

Brochu, P., D. A. Green, T. Lemieux, and J. Townsend (2023). The Minimum Wage, Turnover, and the Shape of the Wage Distribution. *IZA Discussion Paper Series No. 16514*.

Brüll, E. and C. Gathmann (2020). Evolution of the East German Wage Structure. *ZEW Discussion Papers No. 20-081*.

Bruns, B. (2019). Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap. *American Economic Journal: Applied Economics 11*(2), 74–113.

Buchinsky, M. (1994). Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica 62*(2), 405–458.

Bundesinstitut für Bau-, Stadt- und Raumforschung (2019). Indikatoren und Karten zur Raum- und Stadtentwicklung. Accessed online (14/01/2021): https://www.inkar.de/documents/Erlaeuterungen%20Raumbezuege19.pdf.

Bundesministerium für Wirtschaft und Klimaschutz (2023). Öffentliche Konsultation zur Bindung der Auftragsvergabe des Bundes an die Einhaltung von Tarifverträgen ("Bundes-Tariftreue"). Accessed online (17/05/2024): https://www.bmwk.de/Redaktion/DE/Artikel/Service/Gesetzesvorhaben/offentliche-konsultation-zur-bundes-tariftreue.html.

Burauel, P., M. Caliendo, M. M. Grabka, C. Obst, M. Preuss, and C. Schröder (2020). The Impact of the Minimum Wage on Working Hours. *Journal of Economics and Statistics 240*(2-3), 233–267.

Burauel, P., M. Caliendo, M. M. Grabka, C. Obst, M. Preuss, C. Schröder, and C. Shupe (2019). The Impact of the German Minimum Wage on Individual Wages and Monthly Earnings. *Journal of Economics and Statistics 240*(2-3), 201–231.

Caliendo, M., A. Fedorets, M. Preuss, C. Schröder, and L. Wittbrodt (2018). The Short-Run Employment Effects of the German Minimum Wage Reform. *Labour Economics 53*, 46–62.

Caliendo, M., A. Fedorets, M. Preuss, C. Schröder, and L. Wittbrodt (2023). The Short- and Medium-Term Distributional Effects of the German Minimum Wage Reform. *Empirical Economics 64*, 1149–1175.

Caliendo, M., C. Schröder, and L. Wittbrodt (2019). The Causal Effects of the Minimum Wage Introduction in Germany – An Overview. *German Economic Review 20*, 257–292.

Callaway, B. and T. Li (2019). Quantile Treatment Effects in Difference in Differences Models with Panel Data. *Quantitative Economics 10*(4), 1579–1618.

Callaway, B., T. Li, and T. Oka (2018). Quantile Treatment Effects in Difference in Differences Models Under Dependence Restrictions and with Only Two Time Periods. *Journal of Econometrics 206*(2), 395–413.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics 225*(2), 200–230.

Card, D. (1992). Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage. *ILR Review 46*(1), 22–37.

Card, D. (1996). The Effect of Unions on the Structure of Wages: A Longitudinal Analysis. *Econometrica 64*(4), 957–979.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics 128*(3), 967–1015.

Card, D. and A. B. Krueger (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review 84*(4), 772–793.

Card, D. and T. Lemieux (1996). Wage Dispersion, Returns to Skill, and Black-White Wage Differentials. *Journal of Econometrics 74*(2), 319–361.

Card, D., T. Lemieux, and W. C. Riddell (2003). Unions and the Wage Structure. In *International Handbook of Trade Unions*, pp. 246–292. Cheltenham, UK: Edward Elgar Publishing.

Card, D., T. Lemieux, and W. C. Riddell (2004). Unions and Wage Inequality. *Journal of Labor Research 25*(4), 519–562.

Card, D., T. Lemieux, and W. C. Riddell (2020). Unions and Wage Inequality: The Roles of Gender, Skill and Public Sector Employment. *Canadian Journal of Economics/Revue canadienne d'économique 53*(1), 140–173.

Cattaneo, M. D. (2010). Efficient Semiparametric Estimation of Multi-Valued Treatment Effects Under Ignorability. *Journal of Econometrics 155*(2), 138–154.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The Effect of Minimum Wages on Low-Wage Jobs. *Quarterly Journal of Economics 134*(3), 1405–1454.

Chen, L.-Y. and J. Szroeter (2014). Testing Multiple Inequality Hypotheses: A Smoothed Indicator Approach. *Journal of Econometrics 178*, 678–693.

Chernozhukov, V., I. Fernández-Val, S. Han, and A. Kowalski (2019). Censored Quantile Instrumental-Variable Estimation with Stata. *The Stata Journal 19*(4), 768–781.

Chernozhukov, V., I. Fernández-Val, and A. E. Kowalski (2015). Quantile Regression with Censoring and Endogeneity. *Journal of Econometrics 186*(1), 201–221.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2013). Inference on Counterfactual Distributions. *Econometrica 81*(6), 2205–2268.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2020). Quantile and Distribution Regression in Stata: Algorithms, Pointwise and Functional Inference. *Working Paper*.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2022). Fast Algorithms for the Quantile Regression Process. *Empirical Economics 62*, 7–33.

Chernozhukov, V. and H. Hong (2002). Three-Step Censored Quantile Regression and Extramarital Affairs. *Journal of the American Statistical Association 97*(459), 872–882.

Chetverikov, D., B. Larsen, and C. Palmer (2016). IV Quantile Regression for Group-Level Treatments, with an Application to the Distributional Effects of Trade. *Econometrica 84*(2), 809–833.

Collins, W. J. and G. T. Niemesh (2019). Unions and the Great Compression of Wage Inequality in the US at Mid-Century: Evidence from Local Labour Markets. *The Economic History Review 72*(2), 691–715.

Cribb, J., G. Giupponi, R. Joyce, A. Lindner, T. Waters, T. Wernham, and X. Xu (2021). The Distributional and Employment Impacts of Nationwide Minimum Wage Changes. *Institute for Fiscal Studies Working Paper 21*(48).

Datta, N. (2021). Local Monopsony Power. *UniCredit Foundation Working Paper Series No. 161*.

Dauth, W. and J. Eppelsheimer (2020). Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide. *Journal for Labour Market Research 54*(1), 1–14.

Dauth, W., S. Findeisen, and J. Suedekum (2014). The Rise of the East and the Far East: German Labor Markets and Trade Integration. *Journal of the European Economic Association 12*(6), 1643–1675.

Dauth, W., S. Findeisen, and J. Suedekum (2021). Adjusting to Globalization in Germany. *Journal of Labor Economics 39*(1), 263–302.

De Chaisemartin, C. and X. d'Haultfœuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Dickens, R. and A. Manning (2004). Spikes and Spill-Overs: The Impact of the National Minimum Wage on the Wage Distribution in a Low-Wage Sector. *The Economic Journal 114*(494), C95–C101.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica 64*(5), 1001–1044.

DiNardo, J. and D. S. Lee (2004). Economic Impacts of New Unionization on Private Sector Employers: 1984–2001. *Quarterly Journal of Economics 119*(4), 1383–1441.

DiNardo, J. and T. Lemieux (1997). Diverging Male Wage Inequality in the United States and Canada, 1981–1988: Do Institutions Explain the Difference? *ILR Review 50*(4), 629–651.

Dobkin, C., A. Finkelstein, R. Kluender, and M. J. Notowidigdo (2018). The Economic Consequences of Hospital Admissions. *American Economic Review 108*(2), 308–52.

Drechsel-Grau, M., A. Peichl, K. D. Schmid, J. F. Schmieder, H. Walz, and S. Wolter (2022). Inequality and Income Dynamics in Germany. *Quantitative Economics 13*(4), 1593–1635.

Dribbusch, H., S. Lehndorff, and T. Schulten (2017). Two Worlds of Unionism? German Manufacturing and Service Unions Since the Great Recession. *Rough Waters – European Trade Unions in a Time of Crises, Brüssel*, 197–220.

Dube, A. (2019a). Impacts of Minimum Wages: Review of the International Evidence. *Independent Report. UK Government Publication*.

Dube, A. (2019b). Minimum Wages and the Distribution of Family Incomes. *American Economic Journal: Applied Economics 11*(4), 268–304.

Dube, A., L. Giuliano, and J. Leonard (2019). Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior. *American Economic Review 109*(2), 620–63.

Dustmann, C., B. Fitzenberger, U. Schönberg, and A. Spitz-Oener (2014). From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy. *Journal of Economic Perspectives 28*(1), 167–88.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. vom Berge (2022). Reallocation Effects of the Minimum Wage. *Quarterly Journal of Economics 137*(1), 267–328.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009). Revisiting the German Wage Structure. *Quarterly Journal of Economics 124*(2), 843–881.

Ellguth, P. and S. Kohaut (2019). A Note on the Decline of Collective Bargaining Coverage: The Role of Structural Change. *Jahrbücher für Nationalökonomie und Statistik 239*(1), 39–66.

Fan, Y. and Z. Yu (2012). Partial Identification of Distributional and Quantile Treatment Effects in Difference-in-Differences Models. *Economics Letters 115*(3), 511–515.

Farber, H. S., D. Herbst, I. Kuziemko, and S. Naidu (2021). Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data. *Quarterly Journal of Economics 136*(3), 1325–1385.

Felbermayr, G., M. Battisti, and S. Lehwald (2016). Einkommensungleichheit in Deutschland, Teil 1: Gibt es eine Trendumkehr? *ifo Schnelldienst 69*(13), 28–37.

Felgueroso, F., M. J. Pérez-Villadóniga, and J. Prieto-Rodriguez (2008). The Effect of the Collective Bargaining Level on the Gender Wage Gap: Evidence From Spain. *The Manchester School 76*(3), 301–319.

Fernández-Val, I., J. Meier, A. van Vuuren, and F. Vella (2024). Distribution Regression Difference-in-Differences. *Paper presented at the Conference of the International Association for Applied Econometrics 2024*.

Firpo, S. (2007). Efficient Semiparametric Estimation of Quantile Treatment Effects. *Econometrica 75*(1), 259–276.

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional Quantile Regressions. *Econometrica 77*(3), 953–973.

Firpo, S. and C. Pinto (2016). Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures. *Journal of Applied Econometrics 31*(3), 457–486.

Firpo, S. P., N. M. Fortin, and T. Lemieux (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics 6*(2), 1–40.

Fitzenberger, B. (1997). Computational Aspects of Censored Quantile Regression. *Lecture Notes-Monograph Series*, 171–186.

Fitzenberger, B. and A. Doerr (2016). Konzeptionelle Lehren aus der ersten Evaluationsrunde der Branchenmindestlöhne in Deutschland. *Journal for Labour Market Research 49*, 329–347.

Fitzenberger, B., K. Kohn, and A. C. Lembcke (2013). Union Density and Varieties of Coverage: The Anatomy of Union Wage Effects in Germany. *ILR Review 66*(1), 169–197.

Fitzenberger, B., A. Osikominu, and R. Völter (2006). Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Journal of Contextual Economics–Schmollers Jahrbuch* (3), 405–436.

Fitzenberger, B. and A. Seidlitz (2020). Wage Inequality among the Full-Time Employed in Germany Revisited: An Overview. *AStA Wirtschafts-und Sozialstatistisches Archiv 14*, 125–143.

Fitzenberger, B. and K. Sommerfeld (2016). A Sequential Decomposition of the Drop in Collective Bargaining Coverage. *Journal of Economics and Statistics 236*(1), 37–69.

Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder (2019). Metadatenreport. Teil I: Allgemeine und methodische Informationen zur Verdienststrukturerhebung 2014 (EVAS-Nummer: 62111). Version 2. FDZ Standort Hessen.

Fortin, N. M., T. Lemieux, and S. Firpo (2011). Decomposition Methods in Economics. In *Handbook of Labor Economics*, Volume 4, pp. 1–102. Elsevier.

Fortin, N. M., T. Lemieux, and N. Lloyd (2021). Labor Market Institutions and the Distribution of Wages: The Role of Spillover Effects. *Journal of Labor Economics 39*(S2), S369–S412.

Freeman, R. B. (1980). Unionism and the Dispersion of Wages. *ILR Review 34*(1), 3–23.

Freeman, R. B. (1991). How Much Has De-unionisation Contributed to the Rise in Male Earnings Inequality? *NBER Working Paper Series 3826*.

Freyaldenhoven, S., C. Hansen, J. P. Pérez, and J. M. Shapiro (2021). Visualization, Identification, and Estimation in the Linear Panel Event-Study Design. In *Advances in Economics and Econometrics: Twelfth World Congress*. Cambridge: Cambridge University Press. Forthcoming. Draft available at https://scholar.harvard.edu/sites/scholar.harvard.edu/files/shapiro/files/eventstudy.pdf.

Friedrich, M. (2020). Using Occupations to Evaluate the Employment Effects of the German Minimum Wage. *Journal of Economics and Statistics 240*(2-3), 269–294.

Frölich, M. and B. Melly (2013). Unconditional Quantile Treatment Effects Under Endogeneity. *Journal of Business & Economic Statistics 31*(3), 346–357.

Garloff, A. (2019). Did the German Minimum Wage Reform Influence (Un)Employment Growth in 2015? Evidence From Regional Data. *German Economic Review 20*(3), 356–381.

Gartner, H. (2005). The Imputation of Wages Above the Contribution Limit with the German IAB Employment Sample. *FDZ Methodenreport 2*(2005), 1–8.

Giupponi, G., R. Joyce, A. Lindner, T. Waters, T. Wernham, and X. Xu (2024). The Employment and Distributional Impacts of Nationwide Minimum Wage Changes. *Journal of Labor Economics 42*(S1), S293–S333.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics 225*(2), 254–277.

Gopalan, R., B. H. Hamilton, A. Kalda, and D. Sovich (2021). State Minimum Wages, Employment, and Wage Spillovers: Evidence from Administrative Payroll Data. *Journal of Labor Economics 39*(3), 673–707.

Green, D. A. (2014). What Is a Minimum Wage For? Empirical Results and Theories of Justice. *Canadian Public Policy 40*(4), 293–314.

Gregory, T. and U. Zierahn (2022). When the Minimum Wage Really Bites Hard: The Negative Spillover Effect on High-Skilled Workers. *Journal of Public Economics 206*, 1–32.

Göbel, H. (2024). Mindestlohnübermut. *Frankfurter Allgemeine Zeitung, Nr. 112, 15.05.2024*, S. 15.

Hampel, F. R. (1974). The Influence Curve and Its Role in Robust Estimation. *Journal of the American Statistical Association 69*(346), 383–393.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons.

Havnes, T. and M. Mogstad (2015). Is Universal Child Care Leveling the Playing Field? *Journal of Public Economics 127*, 100–114.

Hirsch, B., P. Lentge, and C. Schnabel (2022). Uncovered Workers in Plants Covered by Collective Bargaining: Who Are They and How Do They Fare? *British Journal of Industrial Relations 60*(4), 929–945.

Hoynes, H. W. and A. J. Patel (2018). Effective Policy for Reducing Poverty and Inequality? The Earned Income Tax Credit and the Distribution of Income. *Journal of Human Resources 53*(4), 859–890.

Hoynes, H. W. and A. J. Patel (2022). Replication Data For: Hoynes and Patel "Effective Policy for Reducing Inequality? The Earned Income Tax Credit and the Distribution of Income", *Journal of Human Resources 53(4)*, 859–890, 2018.

Hoynes, H. W. and J. Rothstein (2016). Tax Policy Toward Low-Income Families. *NBER Working Paper Series 22080*.

Jäger, S., S. Noy, and B. Schoefer (2022). The German Model of Industrial Relations: Balancing Flexibility and Collective Action. *Journal of Economic Perspectives 36*(4), 53–80.

Jäger, S., B. Schoefer, and J. Heining (2021). Labor in the Boardroom. *Quarterly Journal of Economics 136*(2), 669–725.

Kim, D. and J. M. Wooldridge (2024). Difference-in-Difference Estimator of Quantile Treatment Effect on the Treated. *Working Paper*.

Knabe, A., R. Schöb, and M. Thum (2014). Der flächendeckende Mindestlohn. *Perspektiven der Wirtschaftspolitik 15*(2), 133–157.

Koenker, R. and G. Bassett Jr. (1978). Regression Quantiles. *Econometrica 46*(1), 33–50.

Kottelenberg, M. J. and S. F. Lehrer (2017). Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Child Care. *Journal of Labor Economics 35*(3), 609–653.

Lee, D. S. (1999). Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage? *Quarterly Journal of Economics 114*(3), 977–1023.

Machado, J. A. F. and J. Mata (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics 20*(4), 445–465.

Manning, A. (2003). *Monopsony in Motion*. Princeton University Press.

Manning, A. (2021). Monopsony in Labor Markets: A Review. *ILR Review 74*(1), 3–26.

Mindestlohnkommission (2020). Dritter Bericht zu den Auswirkungen des gesetzlichen Mindestlohns. Bericht der Mindestlohnkommission an die Bundesregierung nach §9 Abs. 4 Mindestlohngesetz.

Müller, K.-U. and V. Steiner (2011). Beschäftigungswirkungen von Lohnsubventionen und Mindestlöhnen – Zur Reform des Niedriglohnsektors in Deutschland. *Zeitschrift für ArbeitsmarktForschung 44*(1), 181–195.

Nahles, A. (2014). Rede der Bundesministerin für Arbeit und Soziales, Andrea Nahles, zum Entwurf eines Gesetzes zur Stärkung der Tarifautonomie vor dem Deutschen Bundestag am 5. Juni 2014 in Berlin. *Bulletin der Bundesregierung 66*(1), 1–5.

Neumark, D., M. Schweitzer, and W. Wascher (2004). Minimum Wage Effects throughout the Wage Distribution. *Journal of Human Resources 39*(2), 425–450.

Neumark, D. and P. Shirley (2022). Myth or Measurement: What Does the New Minimum Wage Research Say about Minimum Wages and Job Loss in the United States? *Industrial Relations: A Journal of Economy and Society 61*(4), 384–417.

Neumark, D. and W. L. Wascher (2008). *Minimum Wages*. MIT press.

Nichols, A. and J. Rothstein (2016). The Earned Income Tax Credit. In *Economics of Means-Tested Transfer Programs in the United States*, Volume I, pp. 137–218. University of Chicago Press.

Oberfichtner, M. and C. Schnabel (2019). The German Model of Industrial Relations: (Where) Does It Still Exist? *Jahrbücher für Nationalökonomie und Statistik 239*(1), 5–37.

Oberfichtner, M., C. Schnabel, and M. Töpfer (2020). Do Unions and Works Councils Really Dampen the Gender Pay Gap? Discordant Evidence From Germany. *Economics Letters 196*, 1–4.

Oka, T. and K. Yamada (2023). Heterogeneous Impact of the Minimum Wage: Implications for Changes in Between- and Within-Group Inequality. *Journal of Human Resources 58*(1), 335–362.

Plotnick, R. D. (1982). Trends in Male Earnings Inequality. *Southern Economic Journal 48*(3), 724–732.

Powell, D. (2020). Quantile Treatment Effects in the Presence of Covariates. *The Review of Economics and Statistics 102*(5), 994–1005.

Powell, J. L. (1984). Least Absolute Deviations Estimation for the Censored Regression Model. *Journal of Econometrics 25*(3), 303–325.

Powell, J. L. (1986). Censored Regression Quantiles. *Journal of Econometrics 32*(1), 143–155.

Pusch, T. (2024). Effects of the German Minimum Wage on Wages and Household Income. *Journal of Economics and Statistics*, 1–31.

Ransom, T. (2022). Labor Market Frictions and Moving Costs of the Employed and Unemployed. *Journal of Human Resources 57*(S), S137–S166.

Rios-Avila, F. (2020). Statalist Thread on Estimating Quantile Treatment Effects with DiD. Accessed online (02/12/2022): https://www.statalist.org/forums/forum/General-Stata-Discussion/general/1562180-Estimating-a-Did-Quantile-Treatment-Effects.

Rios-Avila, F. and J. de New (2022). Marginal Unit Interpretation of Unconditional Quantile Regression and Recentered Influence Functions Using Centred Regression. *Melbourne Institute Working Paper 14/22*.

Rios-Avila, F. and M. L. Maroto (2024). Moving Beyond Linear Regression: Implementing and Interpreting Quantile Regression Models with Fixed Effects. *Sociological Methods & Research 53*(2), 639–682.

Rossbach, H. (2024). Mindestlohn – Der wahre Tabubruch. *Süddeutsche Zeitung, Nr. 113, 17.05.2024*, S. 4.

Roth, J. and P. H. Sant'Anna (2023). When Is Parallel Trends Sensitive to Functional Form? *Econometrica 91*(2), 737–747.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2023). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *Journal of Econometrics 235*(2), 2218–2244.

Rothe, C. (2010). Nonparametric Estimation of Distributional Policy Effects. *Journal of Econometrics 155*(1), 56–70.

Rothe, C. (2012). Partial Distributional Policy Effects. *Econometrica 80*(5), 2269–2301.

Ruf, K., L. Schmidtlein, S. Seth, H. Stüber, and M. Umkehrer (2021). Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993-2019. *FDZ-Datenreport 3*(2021).

Saint-Paul, G. (2000). *The Political Economy of Labour Market Institutions*. OUP Oxford.

Salverda, W. and D. Checchi (2015). Labor Market Institutions and the Dispersion of Wage Earnings. In *Handbook of Income Distribution*, Volume 2, pp. 1535–1727.

Słoczyński, T. (2022). Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights. *The Review of Economics and Statistics 104*(3), 501–509.

Stansbury, A. and L. H. Summers (2020). The Declining Worker Power Hypothesis: An Explanation for the Recent Evolution of the American Economy. *NBER Working Paper Series 27193*.

Statistisches Bundesamt (2008). Klassifikation Der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). Accessed online (19/01/2021): https://www.destatis.de/static/DE/dokumente/Klassifikation-Wz-2008-3100100089004.pdf.

Statistisches Bundesamt (2016). Qualitätsbericht Verdienststrukturerhebung 2014. Accessed online (25/01/2021): https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verdienste/Verdienststrukturerhebung-2014.pdf?__blob=publicationFile.

Statistisches Bundesamt (2020). Qualitätsbericht Verdienststrukturerhebung 2018. Accessed online (25/01/2021): https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verdienste/Verdienststrukturerhebung-2018.pdf?__blob=publicationFile.

Stewart, M. B. and J. K. Swaffield (2008). The Other Margin: Do Minimum Wages Cause Working Hours Adjustments for Low-Wage Workers? *Economica 75*(297), 148–167.

Tax Policy Center (2023). EITC Parameters. Accessed online (23/03/2023): https://www.taxpolicycenter.org/statistics/EITC-Parameters.

Thomsen, U., J. Ludsteck, and A. Schmucker (2018). Skilled or Unskilled – Improving the Information on Qualification for Employee Data in the IAB Employee Biography. *FDZ-Methodenreport 9*(2018).

Tillé, Y. and M. Langel (2012). Histogram-Based Interpolation of the Lorenz Curve and Gini Index for Grouped Data. *American Statistician 66*(4), 225–231.

Van Kerm, P. (2022). Exchangeably Weighted Bootstrap Schemes. Swiss Stata Conference 2022 10, Stata Users Group.

vom Berge, P., M. Umkehrer, and S. Wager (2023). A Correction Procedure for the Working Hours Variable in the IAB Employee History. *Journal for Labour Market Research 57*(10), 9–16.

vom Berge, P., H. Verbeek, M. Umkehrer, M. Fertig, and S. Bender (2014). Vorbereitende Forschung für die zweite Evaluationsrunde Mindestlöhne: Erschließung neuer Datenquellen. *FDZ-Methodenreporte 03/2014*.

Weesie, J. (2000). Seemlingly Unrelated Estimation and the Cluster-Adjusted Sandwich Estimator. *Stata Technical Bulletin 9*(52).

Xu, R., J. Gao, T. Oka, and Y.-J. Whang (2023). Estimation of Heterogeneous Treatment Effects Using Quantile Regression with Interactive Fixed Effects. *Working Paper*.