

Unintended Consequences and Spill-over Effects of
Family Policies:
Six Essays in Labour and Family Economics

INAUGURAL-DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
doctor rerum politicarum
(Dr. rer. pol.)

am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin



vorgelegt von
Jonas Jessen (M.Sc.)

Berlin, 2021

Erstgutachterin: Prof. Dr. C. Katharina Spieß
Freie Universität Berlin und DIW Berlin

Zweitgutachterin: Prof. Almudena Sevilla, PhD
University College London und IZA

Disputation: November 15, 2021

Acknowledgments

First and foremost, I want to thank my first supervisor (and co-author) C. Katharina Spieß for her continuous support, encouragement, and guidance, and for creating such a positive environment in the department. I am also grateful to Almudena Sevilla for being my second supervisor and for providing valuable feedback for my work. Jochen Kluge was an important mentor and sparked in me an interest in research years before starting my PhD.

I cooperated with several co-authors (Jan Berkes, Mathias Huebener, Robin Jessen, Jochen Kluge, Daniel Kuehnle, Michael Oberfichtner, Sophia Schmitz, C. Katharina Spieß, Sevrin Waights, and Felix Weinhardt) during my PhD and I benefited tremendously from working together. Not all projects came together easily, but through many fruitful discussions, a lot of hard work and creativity, we produced papers I am proud of. Needless to say, it was also fun to work together on a personal level.

In 2016, we jointly started our PhD as the GC cohort. The times shared studying in the big office, our dinner events, the lunches, and other activities made the non-research related aspects of the PhD a great experience. I wrote my PhD at the Education and Family Department of the DIW Berlin. The supportive and friendly atmosphere at my department made the time enjoyable and I am grateful for having had such amazing colleagues. I would like to particularly single out Ludovica Gambaro and Jan Marcus for constant academic guidance despite not (yet) being co-authors. My PhD was partly funded through a scholarship by the German Academic Scholarship Foundation. I also received funding from the Ministry for Education and Research through the Bremen Initiative to Foster Early Childhood Development (BRISE, Förderkennzeichen 01NV1601B). The financial support is much appreciated.

I shared my office briefly with Juliane Stahl and Andreas Leibing, and for about one year each with Georg Camehl and Sophia Schmitz. Thanks for the great atmosphere. For almost two years I shared the office with Jan Berkes (due to COVID-19 we could only properly share it for one year though). Over the last years, we became co-authors, plant meme creators and close friends, and I am grateful for having been on this journey with you.

Last but not least, I want to say a big thank you to my family and friends who have constantly supported me, I couldn't have done it without you. While I cannot list everyone, I particularly would like to thank my parents (Hauke and Di), my brother (Robin) and, above all, Laura for all the love and support.

Berlin, November 29, 2021

Jonas Jessen

Contents

Acknowledgments	III
Erklärung zu Ko-Autorenschaften	VIII
Liste der Vorveröffentlichungen	IX
Rechtliche Erklärung	XI
Abstract	XII
Zusammenfassung	XV
List of Figures	XVIII
List of Tables	XXI
1 Introduction	1
1.1 Motivation	1
1.2 Overview and summary	4
1.3 Contributions	9
2 Punishing Potential Mothers? Evidence for Statistical Employer Discrimination From a Natural Experiment	14
2.1 Introduction	14
2.2 Institutional background	17
2.3 Mechanisms and expected effects	19
2.4 Data and empirical approach	21
2.5 Results	26
2.6 Conclusion	31
3 A Firm-Side Perspective on Parental Leave	33
3.1 Introduction	33

3.2	Background and Institutional Environment	37
3.2.1	Relevant Aspects of the German labour market	37
3.2.2	Family Policies Supporting Women in the Labour Market	38
3.3	Data	41
3.3.1	Data Source: Social Security Records	41
3.3.2	Operationalising Internal and External Substitutes	42
3.3.3	Sample Selection and Treatment Assignment	43
3.3.4	Outcome Variables	46
3.4	Firms and Parental Leave Absences	47
3.4.1	Replacement Hirings and Separations	47
3.4.2	Mothers' Return Behaviour with Short Paid Parental Leave Benefits	50
3.4.3	Mothers' Return Behaviour with Extended Paid Parental Leave	52
3.5	Effects of Extended Parental Leave Absences on Mothers and Firms	54
3.5.1	Empirical Strategy	54
3.5.2	Worker Absences and Effects on Firms	56
3.6	Effects of Extended Parental Leave on Hiring Decisions	61
3.7	Conclusion	64
3.A	Appendix	65
3.A.1	Figures	65
3.A.2	Tables	70
4	Understanding Day Care Enrolment Gaps	76
4.1	Introduction	77
4.2	Institutional background	80
4.3	Data	81
4.3.1	Day care demand	82
4.3.2	Available slots	83
4.3.3	Family background	84
4.4	Day care enrolment gaps and their determinants	85
4.4.1	Enrolment gaps and reasons for non-enrolment	85
4.4.2	Decomposition analysis	89
4.5	The impacts of access barriers on enrolment gaps	91
4.5.1	Local day care shortages	91
4.5.2	Day care fees	94
4.6	Interpretation of the results and policy suggestions	98
4.7	Conclusion	101

Appendix	102
4.A Appendix	102
4.A.1 Interpretation of the parental demand variable	102
4.A.2 Comparison of analysis sample with dropped observations	104
4.A.3 Non-linear estimation	106
4.A.4 Evidence on discrimination in Germany	108
4.A.5 Cost-effectiveness calculations	109
4.A.6 Figures	110
4.A.7 Tables	114
5 Center-Based Care and Parenting Activities	120
5.1 Introduction	120
5.2 Institutional background	123
5.3 Adjustment mechanisms	126
5.4 Data and empirical approach	127
5.4.1 German Time-Use Survey	127
5.4.2 The German Family Panel	130
5.4.3 Empirical approach	132
5.5 Results	133
5.5.1 Effects of center-based care usage	133
5.5.2 Direct and indirect effects	135
5.5.3 Effects of full-day vs. half-day center-based care	137
5.6 Conclusion	141
Appendix	143
5.A Appendix	143
5.A.1 Stylized examples of adjustment mechanisms	143
5.A.2 Data	145
5.A.2.1 Additional descriptive material for time-use data and pairfam	145
5.A.2.2 Activities categories in time-use data	148
5.A.2.3 Data on informal care	151
5.A.3 Center-based care activities with children	153
5.A.4 Heterogeneous effects	155
5.A.5 Tests of sample restrictions	157
5.A.6 Further result tables	159

6	When equality doesn't make it indoors: Gender differences in market and non-market work in the GDR and reunified Germany	166
6.1	Introduction	166
6.2	Institutional setting	171
6.3	Data	172
6.4	Results	173
6.4.1	Heterogeneity by household type	175
6.4.2	Decomposition of the gender gap in housework	178
6.4.3	The role of children	180
6.5	Conclusion	181
6.A	Appendix	183
6.A.1	Figures	183
6.A.2	Tables	184
7	Culture, Children and Couple Gender Inequality	188
7.1	Introduction	188
7.2	German division and reunification	193
7.3	Data and empirical approach	195
7.3.1	Sample criteria and outcomes	196
7.3.2	Additional sources	198
7.3.3	Empirical approach	200
7.4	Results	202
7.4.1	Labour market and domestic outcomes - event study	202
7.4.2	Time-use evidence	206
7.4.3	Attitudes	209
7.5	Robustness	213
7.6	Conclusion	217
7.A	Appendix	219
7.A.1	Figures	219
7.A.2	Tables	229
7.A.3	Robustness	232
7.A.4	Comparison of time-use data from diary data and SOEP	239
8	Conclusion	242
	Bibliography	247

Erklärung zu Ko-Autorenschaften

Diese Dissertation besteht aus sechs (Arbeits-)Papieren, von denen fünf in Zusammenarbeit mit Ko-AutorInnen entstanden sind:

- Jonas Jessen, Robin Jessen und Jochen Kluge: *“Punishing Potential Mothers? Evidence for Statistical Employer Discrimination From a Natural Experiment”*
- Mathias Huebener, Jonas Jessen, Daniel Kuehnle und Michael Oberfichtner: *“A Firm-Side Perspective on Parental Leave”*
- Jonas Jessen, Sophia Schmitz und Sevrin Waights: *“Understanding Day Care Enrolment Gaps”*
- Jonas Jessen, C. Katharina Spiess und Sevrin Waights: *“Center-based Care and Parenting Activities”*
- Jan Berkes, Jonas Jessen und Felix Weinhardt: *“When Equality Doesn’t Make it Indoors: Gender Differences in Market and Non-market Work in the GDR and Reunified Germany”*
- Jonas Jessen: *“Culture, Children and Couple Gender Inequality”*

Liste der Vorveröffentlichungen

- **Publikationen in referierten Fachzeitschriften**

- **Vorpublikation von Kapitel 2:**

- * Jessen, J., Jessen, R., and Kluge, J. (2019). Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment. *Labour Economics*, 59, 164-172.

- <https://doi.org/10.1016/j.labeco.2019.04.002>

- **Vorpublikation von Kapitel 4:**

- * Jessen, J., Schmitz, S., and Waights, S. (2020). Understanding day care enrolment gaps. *Journal of Public Economics*, 190, 104252.

- <https://doi.org/10.1016/j.jpubeco.2020.104252>

- **Beiträge zur Politikberatung**

- **Teilweise basierend auf Kapitel 4 ist erschienen:**

- * Jessen, J., Spieß, C. K., Waights, S., and Judy, A. (2020). Gründe für unterschiedliche Kita-Nutzung von Kindern unter drei Jahren sind vielfältig. *DIW Wochenbericht*, 87(14), 268-275.

- **Arbeitspapiere**

- **Vorpublikation von Kapitel 3:**

- * Huebener, M., Jessen, J., Kuehnle, D. and Oberfichtner, M. (2021). A Firm-Side Perspective on Parental Leave. *IZA DP*, 14478.

- **Vorpublikation von Kapitel 4:**

- * Jessen, J., Schmitz, S. and Waights, S. (2019). Understanding Day Care Enrolment Gaps. *CEP Discussion Paper*, 1650.

- **Vorpublikation von Kapitel 5:**

- * Jessen, J., Spiess, C. K., and Waights, S. (2020). Center-based care and parenting activities. *DIW Discussion Papers*, 1897.

– **Vorpublikation von Kapitel 7:**

- * Jessen, J. (2021). Culture, Children and Couple Gender Inequality. *DIW Discussion Papers*, 1957.

Rechtliche Erklärung

Erklärung gem. §4 Abs. 2 (Promotionsordnung)

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, 29. November 2021

Jonas Jessen

Erklärung gem. §10 Abs. 3 (Promotionsordnung)

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe. Auf dieser Grundlage und in Zusammenarbeit mit meinen Ko-AutorInnen habe ich die Arbeit selbstständig verfasst.

- Software:
 - Stata Versionen 14, 15, 16 und 17
 - QGIS 3.16.1
 - L^AT_EX mit Overleaf
- Literatur: siehe Literaturverzeichnis

Berlin, 29. November 2021

Jonas Jessen

Abstract

This dissertation consists of six independent chapters contributing to the literature of labour and family economics. The main topic concerns how family policies impact on gender and socio-economic inequality in, at times, unintended ways.

Chapter 2 uses administrative linked employer-employee data to examine whether employers statistically discriminated against women of childbearing age (potential mothers) when they incurred direct costs associated with motherhood. Before 2006, large firms in Germany were obliged to pay for the generous maternity protection of female employees, such that firms' expected wage costs depended on employees' gender and age. From 2006 onward, all firms paid for maternity protection by contributing to the statutory health insurance system, where the contribution depends only on the number of employees and their wages and is thus independent of gender and age. We provide evidence that the reform was followed by an increase in female relative wages within large firms. This reform effect provides evidence for statistical employer discrimination in the pre-2006 setup.

Chapter 3 takes a firm-side perspective on parental leave. Motherhood and parental leave interrupt employment relationships, likely imposing costs on firms. We document that mothers who are difficult to replace internally take shorter leave and that their firms hire replacements more often. Introducing more generous parental leave benefits erases the link between mothers' internal replaceability and their leave duration. In firms with few internal substitutes this reduces employment in the short-, but not longer-term. Firms respond by hiring fewer women of childbearing age into occupations where they are difficult to replace internally. Taken together, motherhood and generous parental leave policies burden firms that have few internal substitutes available.

Chapter 4 aims to improve the understanding of day care enrolment gaps by family background in a country with a universal day care system (Germany). Research demonstrates that children of parents with lower educational attainment and children of migrant parents may benefit the most from day care, making it important to understand why such enrolment gaps exist. We use a unique data set that records both parental demand for day care and actual usage to investigate determinants using com-

plementary decomposition and quasi-experimental analyses. Our decomposition shows that (a) differences in demand are important but do not fully explain the enrolment gaps, (b) large shares of the gaps are unexplained, especially for migrant parents, and (c) the heterogeneous effects of access barriers (shortages and fees) may explain some of the remaining gaps. Our quasi-experimental design finds that reducing shortages significantly decreases the enrolment gap by parental education but not by parental migrant status. Similarly, using the synthetic control method we show that a reduction of fees reduces only the gap by parental education. We discuss implications for policy.

Chapter 5 estimates the effect of day care on parenting activities using time-diary and survey data for Germany. This is the first such study in economics to pay careful attention to issues of selection bias, and to provide a conceptual framework of underlying mechanisms. We find that while day care strongly reduces the amount of time parents spent with their child, parenting activities are only reduced by a few minutes per day. During non-center hours, parenting activities are not affected. An analysis of non-parenting activities reveals that day care is used to take up paid work, but also partly to ease time constraints. A reduction of leisure and sleep during non-center hours suggests that an increase in motivation may be responsible for keeping parenting activities constant during those times. Our findings represent novel evidence that activities in the home environment are a complement to day care, highlighting a credible alternative mechanism for child development effects of day care.

Chapter 6 uses novel time-use data from the GDR and reunified Germany and finds that women in the GDR, and later in East Germany, spend more time in paid work and less time doing housework, compared to West Germany. However, decomposing these gender housework gaps between the West and the East, we find that they are similar once individual time constraints are accounted for. Individual housework contributions are shown to be almost orthogonal to the partner's labour supply. We discuss implications for the nature of gender norms, and effects of labour market policy targeted at gender gaps.

Chapter 7 examines how culture impacts within-couple gender inequality. Exploiting the setting of Germany's division and reunification, I compare child penalties of couples socialised in a more gender-egalitarian culture (East Germany) to those in a gender-traditional culture (West Germany). Using a household panel, I show that the long-run child penalty on the female income share is 26.9 percentage points in West German couples, compared to 15.5 in East German couples. Additionally, the arrival of children leads to a stronger increase in the share of housework performed by West German women and they are responsible for a larger share of child care than those from the East. A battery of robustness checks confirms that differences between East and

West socialised couples are not driven by current location, economic factors, day care availability or other smooth regional differences. I add to the main findings by using time-use diary data from the GDR and reunified Germany, comparing parents with childless couples of similar ages. This provides a rare insight into gender inequality in the GDR and allows to compare the effect of children in the GDR to the effects in East and West Germany after reunification. Lastly, I show that attitudes towards maternal employment are more egalitarian among East Germans, but that the arrival of children leads to more traditional attitudes for both East and West Germans. The findings confirm that socialisation has a strong impact on child penalties and thus on gender inequality as a whole.

Zusammenfassung

Diese Dissertation besteht aus sechs unabhängigen Kapiteln, die einen Beitrag zur Literatur im Bereich der Arbeits- und Familienökonomie leisten. Das Hauptthema ist die Frage, wie sich Familienpolitik auf geschlechtsspezifische und sozioökonomische Ungleichheit auswirkt, manchmal auf unbeabsichtigte Weise.

Kapitel 2 verwendet administrative, verknüpfte Arbeitgeber-Arbeitnehmer-Daten, um zu untersuchen, ob Arbeitgeber Frauen im gebärfähigen Alter (potenzielle Mütter) statistisch diskriminierten als ihnen direkte Kosten im Zusammenhang mit der Mutterschaft entstanden. Vor 2006 waren große Firmen in Deutschland verpflichtet, für den großzügigen Mutterschutz ihrer weiblichen Angestellten zu zahlen, sodass die erwarteten Lohnkosten der Firmen vom Geschlecht und Alter der Angestellten abhingen. Ab 2006 zahlten alle Firmen für den Mutterschutz, indem sie einen Beitrag zur gesetzlichen Krankenversicherung leisteten, wobei der Beitrag nur von der Anzahl der Beschäftigten und deren Lohn abhängt und somit unabhängig von Geschlecht und Alter ist. Wir weisen nach, dass die Reform zu einem Anstieg der relativen Löhne von Frauen in größeren Firmen geführt hat. Dieser Reformeffekt liefert Belege für eine statistische Arbeitgeberdiskriminierung in der Zeit vor 2006.

Kapitel 3 nimmt eine Firmenperspektive auf Elternzeit ein. Mutterschaft und Elternzeit unterbrechen Beschäftigungsverhältnisse, was Kosten für Unternehmen verursacht. Wir dokumentieren, dass Mütter, die intern schwer zu ersetzen sind, kürzere Elternzeiten nehmen und dass ihre Firmen häufiger einen Ersatz einstellen. Die Einführung großzügigerer Entgeltersatzleistungen hebt den Zusammenhang zwischen der internen Substituierbarkeit von Müttern und der Länge ihrer Elternzeit auf. In Firmen mit wenigen internen Substituten reduziert dies die Beschäftigung kurz-, aber nicht längerfristig. Die Unternehmen reagieren darauf, indem sie weniger Frauen im gebärfähigen Alter in Berufen einstellen, in denen sie intern schwer zu ersetzen sind. Zusammengenommen belasten Mutterschaft und großzügige Elternzeitregelungen Firmen, die nur wenige interne Substitute zur Verfügung haben.

Kapitel 4 zielt darauf ab, das Verständnis von Unterschieden in Kita-Nutzung nach familiärem Hintergrund in einem Land mit einem universellen Kinderbetreuungssystem

(Deutschland) zu verbessern. Die Forschung zeigt, dass Kinder von Eltern mit niedrigerem Bildungsniveau und Kinder von Eltern mit Migrationshintergrund am meisten von der Kindertagesbetreuung profitieren, weshalb es wichtig ist zu verstehen, warum Unterschiede in der Nutzung bestehen. Wir verwenden einen einzigartigen Datensatz, der sowohl die elterliche Nachfrage nach Betreuung als auch die tatsächliche Nutzung erfasst, um die Determinanten mittels einer Dekomposition und komplementären quasi-experimentellen Analysen zu untersuchen. Unsere Dekomposition zeigt, dass (a) Unterschiede in der Nachfrage wichtig sind, aber die Unterschiede in der Kita-Nutzung nicht vollständig erklären, (b) ein großer Teil der Nutzungsunterschiede unerklärt bleibt, insbesondere bei Eltern mit Migrationshintergrund, und (c) die heterogenen Effekte von Zugangsbarrieren (mangelnde Plätze und Gebühren) einen Teil der verbleibenden Lücken erklären können. Unser quasi-experimentelles Design zeigt, dass die Reduzierung von Engpässen die Unterschiede nach elterlicher Bildung signifikant verringert, nicht aber nach elterlichem Migrationshintergrund. In ähnlicher Weise zeigen wir unter Verwendung der synthetischen Kontrollmethode, dass eine Reduzierung der Gebühren nur die Lücke nach elterlicher Bildung reduziert. Wir diskutieren Implikationen für die Politik.

Kapitel 5 schätzt den Effekt von Kita-Nutzung auf elterliche Aktivitäten unter Verwendung von Zeitverwendungs- und Umfragedaten für Deutschland. Dies ist die erste derartige Studie in den Wirtschaftswissenschaften, die sorgfältig auf Selektion achtet und einen konzeptionellen Rahmen der zugrunde liegenden Mechanismen liefert. Wir zeigen auf, dass die Betreuung in einer Kindertagesstätte zwar die Zeit, die die Eltern mit ihrem Kind verbringen, deutlich reduziert, die interaktiven elterlichen Aktivitäten jedoch nur um wenige Minuten pro Tag verringert werden. Außerhalb der Betreuungszeiten sind die elterlichen Aktivitäten nicht betroffen. Eine Analyse von anderen Aktivitäten zeigt, dass die Betreuung in der Kindertagesstätte genutzt wird, um bezahlte Arbeit aufzunehmen, aber auch teilweise, um zeitliche Restriktionen zu verringern. Die Reduktion von Freizeit und Schlaf außerhalb der Betreuungszeiten deutet darauf hin, dass eine erhöhte Motivation dafür verantwortlich sein könnte, die elterlichen Aktivitäten während dieser Zeiten konstant zu halten. Unsere Ergebnisse liefern Nachweise dafür, dass Aktivitäten in der häuslichen Umgebung ein Komplement zur Kita-Betreuung darstellen und einen glaubwürdigen alternativen Mechanismus für die Auswirkungen der Kita-Betreuung auf die kindliche Entwicklung aufzeigen.

Kapitel 6 verwendet neuartige Zeitverwendungsdaten aus der DDR und dem wiedervereinigten Deutschland und stellt fest, dass Frauen in der DDR und später in Ostdeutschland im Vergleich zu Westdeutschland mehr Zeit mit bezahlter Arbeit und weniger Zeit mit Hausarbeit verbringen. Wenn man jedoch diese geschlechtsspezifischen

Unterschiede bei der Hausarbeit zwischen West und Ost aufschlüsselt, stellt man fest, dass sie ähnlich sind, wenn man die individuellen Zeiteinschränkungen berücksichtigt. Wir zeigen außerdem, dass der individuelle Beitrag zur Hausarbeit nahezu orthogonal zum Arbeitsangebot des Partners ist. Wir diskutieren Implikationen für die Natur von Geschlechternormen und die Auswirkungen von Arbeitsmarktpolitik, die auf Geschlechterunterschiede abzielt.

Kapitel 7 untersucht, wie Kultur sich auf die Ungleichheit zwischen den Geschlechtern innerhalb von Paaren auswirkt. Vor dem Hintergrund der deutschen Teilung und Wiedervereinigung vergleiche ich die Effekte von Kindern auf verschiedene Ergebnisse (*child penalties*) von Paaren, die in einer eher geschlechteregalitären Kultur (Ostdeutschland) sozialisiert wurden, mit denen, die in einer geschlechtertraditionellen Kultur (Westdeutschland) sozialisiert wurden. Anhand eines Haushaltspanels zeige ich, dass die langfristige child penalty auf den weiblichen Einkommensanteil bei westdeutschen Paaren 26,9 Prozentpunkte beträgt, verglichen mit 15,5 bei ostdeutschen Paaren. Ich zeige außerdem, dass Kinder zu einem stärkeren Anstieg des Anteils der Hausarbeit führen, den westdeutsche Frauen leisten, und dass sie für einen größeren Anteil der Kinderbetreuung verantwortlich sind als ostdeutsche Frauen. Eine Reihe von Robustheitstests bestätigt, dass die Unterschiede zwischen ost- und westdeutsch sozialisierten Paaren nicht durch den aktuellen Wohnort, wirtschaftliche Faktoren, die Verfügbarkeit von Kinderbetreuung oder andere regionale Differenzen getrieben werden. Ich ergänze die Hauptergebnisse durch die Verwendung von Zeitverwendungsdaten aus der DDR und dem wiedervereinigten Deutschland, indem ich Eltern mit kinderlosen Paaren ähnlichen Alters vergleiche. Dies bietet einen seltenen Einblick in die Ungleichheit zwischen den Geschlechtern in der DDR und erlaubt es, den Effekt von Kindern in der DDR mit den Effekten in Ost- und Westdeutschland nach der Wiedervereinigung zu vergleichen. Schließlich zeige ich, dass die Einstellungen zur mütterlichen Erwerbstätigkeit bei Ostdeutschen egalitärer sind, dass aber Kinder sowohl bei Ost- als auch bei Westdeutschen zu traditionelleren Einstellungen führen. Die Ergebnisse bestätigen, dass die Sozialisation einen starken Einfluss auf child penalties und damit auf die Ungleichheit der Geschlechter insgesamt hat.

List of Figures

1.1	Connection between chapters	5
2.1	Firms' costs per maternity protection period	18
2.2	Density plot of mothers' age at childbirth	23
2.3	Share of female employees in small and large firms, 2001-2010	27
2.4	Average full-time wages of men and women in large firms (> 30 employees), 2001-2010	28
3.1	Return to pre-birth firm by birth semester	41
3.2	Hirings around childbirth	48
3.3	Hirings by availability of substitutes	49
3.4	Separations by incumbents and replacements	49
3.5	Return to pre-birth firm in setting with low paid leave entitlements	51
3.6	Return to pre-birth firm in setting with extended paid leave entitlements	53
3.7	Event study of parental leave reform effects on mothers' and firms' outcomes	57
3.8	Effect heterogeneity - event study of parental leave reform effects on firm outcomes	60
3.9	Composition of hirings	62
3.10	Trend deviations in hiring composition by demographic group and internal substitutes	63
3.A.1	Return to employment before and after the 2007 parental leave reform	65
3.A.2	Occupations by internal substitutes	66
3.A.3	Size of firms	66
3.A.4	Hirings around childbirth - 2007 births	67
3.A.5	Hirings by availability of substitutes - 2007 births	67
3.A.6	Separations by incumbents and replacements - 2007 births	68
3.A.7	Hirings around childbirth in other occupations	68
3.A.8	Return to pre-birth firm and availability of internal substitutes - discrete categorisation	69

3.A.9	Density test of births	69
4.1	Day care enrolment, demand and shortage by age	84
4.2	Gaps in day care enrolment, demand and rationing probability	87
4.3	The effect of fee elimination in Hamburg on the enrolment gap (by parental education)	96
4.4	The effect of fee elimination in Hamburg on the enrolment gap (by parental migrant status)	97
4.A.1	Parental demand regressed on factors/reasons for not using day care	103
4.A.2	The enrolment gap by parental background at different levels of shortages	107
4.A.3	Shortages in day care	110
4.A.4	Day care enrolment and rationing by birth country of the parents	111
4.A.5	Conditional gaps in day care enrolment, demand and rationing probability	112
4.A.6	Interaction term of Table 4.4 by sample restriction	113
5.1	Share of children enrolled in center-based care by region, age group, and time	125
5.2	Mothers' activities on weekdays by usage of center-based care	131
5.3	Effects by time of day and education	136
5.4	Full-day vs half-day care effects by time of day and education, wave 3 only (2012/13)	139
5.A.1	Adjustment of parenting activities with use of center-based care	143
5.A.2	Parents' activities by usage of center-based care	146
5.A.3	Average time use for mothers and fathers by survey wave	147
5.A.4	Shared activities with the child by half- or full-day usage of center-based care	150
5.A.5	Heterogeneity in overall effects on parenting activities – mothers and fathers pooled	156
5.A.6	Effects of center-based care on parenting activities for boys and girls	156
5.A.7	Comparison of coefficients by sample restriction I	157
5.A.8	Comparison of coefficients by sample restriction II	158
6.1	Female share of housework and paid work	168
6.2	Cumulative distributions of paid work and housework	174
6.3	Female share of housework, female and male minutes of housework by share of paid work in the GDR (1985/90)	176
6.4	Shares of housework by share of paid work in East and West Germany (1991/92)	177

6.A.1	Minutes of housework by gender and share of paid work in East and West Germany (1991/92)	183
7.1	Female income shares in East and West German couples	189
7.2	Female labour force participation	195
7.3	Impact of children on labour market gender inequality	203
7.4	Impact of children on domestic gender inequality	205
7.5	Distribution of activities - time-use data	209
7.6	Maternal employment by child age	211
7.A.1	Observations by event time	219
7.A.2	Specialisation index	220
7.A.3	Average values by event time	221
7.A.4	Impact of children on female income share - by number of children . .	222
7.A.5	Impact of children on gender inequality in the labour market - gaps and discrete outcomes	223
7.A.6	Impact of children on gender inequality in the labour market - individual level	224
7.A.7	Impact of children on domestic gender inequality - weekly	225
7.A.8	Impact of children on total hours of work	226
7.A.9	Impact of children on domestic gender inequality - gaps	226
7.A.10	Impact of children on gender inequality in the labour market - individual level	227
7.A.11	Impact of children on desired working hours - individual level	228
7.A.12	Couple dissolution	228
7.A.13	East-West long-run difference with states excluded	233
7.A.14	Placebo treatment assignment	234
7.A.15	West German counties split by income per capita	235
7.A.16	West German counties split by day care availability	236
7.A.17	German counties and distance to inner border	237
7.A.18	Spatial RD	238

List of Tables

2.1	Summary statistics: firms by size and sector	24
2.2	Estimates of the reform effect on log full-time wages within large firms	29
2.3	Estimates of the reform effect on log full-time wages in large firms—by sector	30
3.1	Summary statistics and balancing	45
3.2	Summary event study estimates - mothers	58
3.3	Summary event study estimates - firm	59
3.A.1	Comparison of mothers in analysis sample with excluded observations	70
3.A.2	Comparison of firms in analysis sample with excluded observations . .	71
3.A.3	Summary statistics by internal substitutes	72
3.A.4	Balancing by internal substitutes (DD coefficients)	73
3.A.5	Relationship between availability of internal substitutes and parental leave length	74
3.A.6	Summary event study estimates: Mothers - employment at any firm .	75
4.1	Summary statistics	86
4.2	Reasons for not using day care	88
4.3	Oaxaca-Blinder decomposition of the enrolment gap	91
4.4	Day care enrolment, family background and regional shortages	93
4.A.1	Comparison of analysis sample with dropped observations	105
4.A.2	Share of children under three years in day care – comparison of KiBS and official statistics	114
4.A.3	Characteristics by country of origin	115
4.A.4	Day care enrolment, family background and regional shortages – alter- native family background definition	116
4.A.5	Day care enrolment, family background and regional shortages – leave- one-out shortage	117
4.A.6	Synthetic Hamburg weights	118
4.A.7	Day care enrolment, family background and fees	119

5.1	Sample characteristics by enrollment in center-based care	130
5.2	Effects of center-based care on parenting activities	134
5.3	The effect of full-day care on parenting and non-parenting activities .	140
5.A.1	Characteristics of pairfam sample	145
5.A.2	Overview of activities in time-use data, 2012/13 wave	148
5.A.3	List of detailed parenting activities in time-use data, 2012/13 wave . .	149
5.A.4	Weekly hours in care - SOEP	151
5.A.5	Usage of formal and informal care	152
5.A.6	Activities in center-based care	154
5.A.7	Effects of center-based care on parents' time spent on parenting and non-parenting activities, by time of day and education	159
5.A.8	Effects of center-based care on parents' time spent on specific parenting activities, by time of day and education	160
5.A.9	Effects of full-day vs. half-day center-based care on parenting and non- parenting activities using one wave (2012/13), by time of day and ed- ucation	161
5.A.10	The effect of full-day care on parenting activities (Oster bounds) . . .	162
5.A.11	The effect of full-day care on non-parenting activities and other parent- and child-related outcomes; pairfam - Oster bounds	163
5.A.12	The effect of full-day care on parenting activities - alternative full-day assignment	165
6.1	Summary statistics of time-use data	173
6.2	Decomposition of the housework gap - East and West Germany (1991/92)	179
6.3	Decomposition of the housework gap by children - East and West Ger- many (1991/92)	180
6.A.1	Countries and surveys used from the Multinational Time Use Study .	184
6.A.2	Time-use data: detailed activities (minutes per weekday)	185
6.A.3	Decomposition of the female housework share - East and West Germany (1991/92)	186
6.A.4	Decomposition of the housework gap (cooking, cleaning, shopping) - East and West Germany (1991/92)	186
6.A.5	Decomposition of the housework gap - East and West Germany (2001/02, 2012/13 and pooled)	187
7.1	Overview of analysis sample	197
7.2	Pre-birth characteristics	199
7.3	Long-run impacts of children	207

7.4	East-West differences in attitudes and the impact of children	212
7.A.1	Long-run impacts of children - full sample and sample with pre- and post-birth observations	229
7.A.2	Test of approximated control group	230
7.A.3	Couple-level sums for main outcomes	230
7.A.4	Impact of children - time-use data	231
7.A.5	Estimates based on current location	232
7.A.6	Both partners born in Germany	232
7.A.7	Time use comparison	241

CHAPTER 1

Introduction

1.1 Motivation

Women's labour force participation and earnings have increased substantially in the last decades, leading to a pronounced reduction in gender inequality across high-income countries. However, gender inequalities have not been fully eradicated and gender gaps in labour force participation and earnings, but also domestic work, have despite this progress remained stubbornly persistent (e.g., Bertrand, 2020; Olivetti and Petrongolo, 2016). Policy interest has increased strongly and research on gender inequality has become a prominent field in economics (Sevilla, 2021). Early work on gender inequality has predominantly focused on human capital and discrimination as key drivers (Altonji and Blank, 1999). The education gap has over time been closed in all high-income countries with women mostly being more likely to hold college degrees than men (Becker et al., 2010), so human capital cannot be a main driver anymore. Since the late 1990s, the role of children for gender inequality has been identified as a key factor (see Lundberg and Rose, 2000; Waldfogel, 1998, for early work on this). A seminal contribution by Kleven et al. (2019b) has documented and quantified that a large share of remaining gender inequality can now be attributed to the arrival of children.¹

Both the reduction of gender inequality as well as its persistence have been analysed through the lens of cultural and institutional factors. Due to the large role that children play in explaining gender inequality, policies aimed at supporting families with young children are of particular interest. Olivetti and Petrongolo (2017) review the literature on economic consequences of family policies and estimate cross-county correlations of

¹This implies by no means that discrimination does not play a role. Kleven et al. (2019b) show for Denmark that residual, i.e., unexplained, gender inequality remains. And, more importantly, their decomposition shows the total effect of children. Both anticipated and realised fertility decisions continue to disproportionately affect women, which means that part of the mechanism of having children may operate through (statistical) discrimination against women.

those policies.² While building on this extensive literature, the first chapters of this dissertation follow a different approach. Rather than focusing on individual labour market consequences of such policies, I and my co-authors analyse important aspects that the literature has despite their importance turned less attention to.

Parental leave policies are a popular policy tool to help parents reconcile work and family life. A plethora of studies has considered the effects of parental leave reforms on mothers' labour market careers, e.g., see Lalive and Zweimüller (2009) and Kleven et al. (2020) for Austria, Schönberg and Ludsteck (2014) and Kluve and Schmitz (2018) for Germany, Ginja et al. (2020a) for Sweden, Dahl et al. (2016) for Norway, or Baum and Ruhm (2016) and Bana et al. (2020) for California. An overall consistent finding of these studies is that parental leave legislation affects employment and earnings in the short-run, but long-run effects tend to be small or zero. But parental leave policies may also affect mothers' or—more generally—young women's labour market careers through other channels than direct labour market effects (as has already been noted in early work by Ondrich et al., 1996).

Those unintended effects of such policies are the topic of **Chapters 2 and 3** of this dissertation. In **Chapter 2** we analyse whether young women of fertile age (henceforth *potential mothers*) were statistically discriminated against by employers when these faced larger direct costs of motherhood by having to pay generous wage replacements during maternity protection. The following **Chapter 3** takes a more direct firm-side perspective on parental leave and analyses how firms are affected by birth-related absences in general and a reform introducing more generous parental leave benefits for middle- and high-income mothers. If firms are harmed by (longer) parental leave absences, then they may try to avoid those absences and adjust their hiring and through this affect young women more broadly.

Another family policy implemented across high-income countries is publicly funded or highly subsidised day care. The early main motivation of expanding day care provision was to facilitate employment of both parents—although providing an educational environment for children was always a consideration as well—so in essence to enable mothers to participate in the labour force as well as these are mostly primary care providers (Kimmel, 1998). Numerous studies have analysed the effects on maternal employment,³ see Bauernschuster and Schlotter (2015); Müller and Wrohlich (2020) for evidence on Germany or, e.g., Baker et al. (2008); Berlinski and Galiani (2007); Cascio

²In a thorough assessment of Austrian parental leave and child care policies, Kleven et al. (2020) use administrative data to analyse their effects on gender inequality over half a century and find that these had *virtually no impact*.

³Another widely studied aspect is the effect of day care provision on fertility (see Bauernschuster et al., 2016; Hank et al., 2004; Rindfuss et al., 2010, for studies on Germany and Norway, respectively).

(2009); Gelbach (2002); Havnes and Mogstad (2011a); Nollenberger and Rodríguez-Planas (2015) for studies from other countries. Another policy objective has become just as central over time; the potential of day care to foster child development and to reduce differences in their development by parental background.

In contrast to targeted programs that are usually aimed at providing care for socially disadvantaged children, most (continental) European countries have universal day care systems. These are meant to ensure that children of all backgrounds have access to day care. Slots are either directly provided by the state or heavily subsidised with mostly income-dependent fees.⁴ Studies from Germany (Cornelissen et al., 2018; Felfe and Lalive, 2018) suggest that children from migrant or lower educated parents benefit disproportionately in their development from attending day care. Studies from other European countries (Drange and Havnes, 2019; Felfe et al., 2015; Datta Gupta and Simonsen, 2010; Havnes and Mogstad, 2011b, 2015) reveal similar patterns.⁵ Yet despite the extensive evidence on (more) beneficial effects for disadvantaged children, in many countries—including Germany—these are often underrepresented in day care (e.g., Schober and Spiess, 2013). I.e., despite the intended universality, parental background still matters. In **Chapter 4** we seek to better understand why those enrolment gaps for children under three persist and combine careful descriptive evidence with a decomposition of gaps and quasi-experimental analyses.

Development effects of day care on children are commonly attributed to the institutional environment at day care centres compared to the counterfactual care environment if children did not attend (Kline and Walters, 2016). More beneficial effects for disadvantaged children are then explained by differences in those environments (Cornelissen et al., 2018). However, the literature rarely addresses (Kuger et al., 2019, being an exception) that usage of day care may have spill-over effects on the home environment. In **Chapter 5** we analyse how parenting activities differ when children are attending day care or not. Rich time-use data give insights into how parents adjust both parenting and other activities, and through this helps to shed light on the mechanisms of development effects of day care on children.

The final two chapters of this dissertation again deal with the topic of gender inequality. In contrast to the first chapters which focused on effects of policies concerning parental leave absences, these final chapters look at gender inequality more broadly. The importance of gender norms is analysed in the context of Germany's division and

⁴In Germany with its progressive fee structure, as of 2015 parents paid on average below 10% of their household net income for day care fees (Schmitz et al., 2017). Since then, many federal states have continued to abolish fees for some or all age groups (Huebener et al., 2020).

⁵In a setting with lower quality of care, Fort et al. (2020) have identified negative development effects of attending day care for children from advantaged backgrounds.

reunification where during the division East and West Germans were exposed to very different policies regarding—among others—maternal labour force participation.

Gender norms have been found to be influential factors in explaining gender inequalities. While it has received little attention by economists for a long time, the role of culture in explaining differences in gender norms has in recent years increasingly taken centre stage (Giuliano, 2021). Over time and through various mechanisms, including observing female employment and social learning, men and women have developed more positive views on maternal employment (Fernández et al., 2004; Fogli and Veldkamp, 2011). Fortin (2005) showed that more egalitarian attitudes are associated with higher female labour force participation rates and lower gender pay gaps.

The *natural experiment* of Germany’s division and reunification has been used extensively by researchers to examine differences in gender norms (e.g., Bauernschuster and Rainer, 2012; Beblo and Görges, 2018; Campa and Serafinelli, 2019; Zoch, 2021), gender inequality in income (Lippmann et al., 2020; Sprengholz et al., 2020), and housework (Cooke, 2007). Consistently, East Germany (formerly the socialist German Democratic Republic, GDR) has been shown to be more gender egalitarian.

In **Chapter 6** we look at how gender inequality in market work is linked to inequality in housework. In the GDR, female and especially maternal employment was strongly promoted by the government and gender differences in labour participation and working hours were small. In the Federal Republic of Germany (FRG), gender policies were more traditional. As a result, female labour force participation in the GDR (and later East Germany) was much higher than in the FRG (West Germany). In this chapter we use time-use data from the GDR and from reunified Germany to analyse whether this also resulted in a more gender egalitarian distribution of household tasks beyond a mechanical effect of women spending less time at home. I.e., do policies that lead to a more even distribution of working hours also spill over to the domestic sphere?

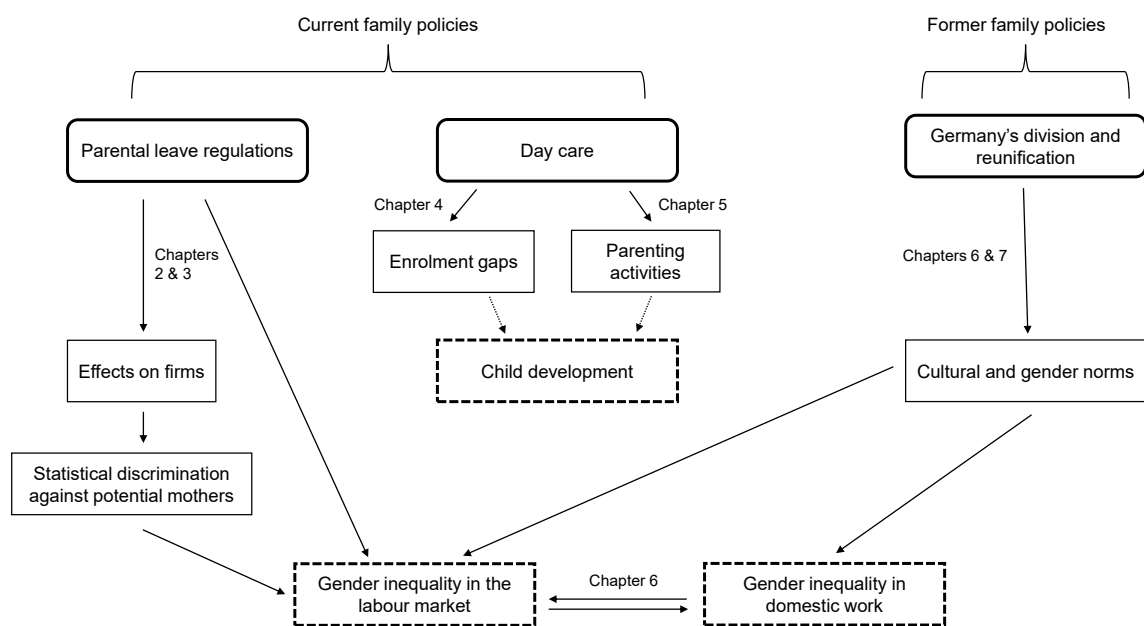
In the final **Chapter 7** of this dissertation I compare child penalties in East and West German couples using two household panels and time-use data. I first show that the female income share is almost similar in East and West German couples and that a large share of widely documented differential gender inequality is due to stronger negative effects of having children in West German couples. Beside labour market inequality, I also consider the effects of children on domestic work and on attitudes.

1.2 Overview and summary

This dissertation comprises six empirical papers in labour and family economics. As individual papers, each chapter is self-contained and constitutes independent contributions to the economic literature. Still, the research questions are closely linked as

outlined in the previous section. The chapters can broadly be categorised into three content-defined groups; **Chapters 2 and 3** concern labour market effects of German parental leave policies at the individual- and firm-level. **Chapters 4 and 5** deal with the German day care system. Finally, **Chapters 6 and 7** look at gender inequality under the lens of Germany's division and reunification. Despite covering different topics, an overarching topic of the chapters of this dissertation is that they examine inequalities—gender and socio-economic—and how family policies have impacted those in, at times, unintended ways.

Figure 1.1: Connection between chapters



Source: Own illustration

Figure 1.1 provides an overview on how the chapters of this dissertation are linked. The bold-framed rounded rectangles at the top depict the aforementioned three groups. The first two groups, *parental leave regulations* and *day care*, relate to current family policies. Parental leave regulations impact female employment directly and also through their effects on firms. Due to statistical discrimination the effects can go beyond those being directly affected by the policies. The other policy group I consider is formal day care. Through its effect on parenting activities and due to enrolment gaps by parental background, day care impacts child development (heterogeneously). During *Germany's division*, maternal (and, more general, female) employment was promoted through various family and other policies in the GDR. In the FRG, in contrast, the system of family policies was gender-traditional and favoured a (main) male breadwinner model. Those policies have impacted cultural and gender norms, which

had contemporaneous and long-lasting effects on gender inequality in the labour market and in domestic work.

Overall, the chapters relate to three main outcomes of interest for public policy, which are shown in bold, dashed rectangles. Two of those, gender inequality in the labour market and in domestic work, are measured and analysed directly. Child development is not considered directly, but it motivates the chapters and conclusions regarding this outcome are drawn from the literature. In the following I provide brief summaries of the chapters of this dissertation.

Chapter 2 investigates statistical discrimination against *potential mothers* in the labour market. Despite its appeal as a theoretical concept (Arrow, 1973; Phelps, 1972), empirical evidence for statistical discrimination against potential mothers in the labour market is scarce. In the analysis we take advantage of a German reform in 2006 as a quasi-experiment to test whether firms were statistically discriminating against potential mothers when the direct cost of maternity leave was higher. Larger firms had to cover a large share of full wage replacements over 14 weeks until this was abolished with the 2006 reform as the Federal Constitutional Court was concerned about statistical discrimination and its detrimental effect on gender equality. Using administrative linked employer-employee data, we compare wage trajectories of men and women in affected firms with difference-in-differences and trend-break models. Our results reveal that compared to the counterfactual reducing the cost of motherhood for firms increases relative female wages by a total of 2.5% over five post-reform years. As predicted, effects are larger in high-wage firms. The findings imply that it is crucial for policy makers to identify regulations that could result in statistical discrimination against mothers and to consider compensations of costs which occur asymmetrically to mothers (or women) only.

In **Chapter 3** we take a direct firm-side perspective on parental leave and analyse in detail how firms are impacted by birth-related absences of mothers. Motherhood and parental leave interrupt employment relationships and thus affect firms directly; firms can either replace the absent mother (internally or externally) or face a reduction in employment and through this, *ceteris paribus*, in value added. As in **Chapter 2**, we build on rich administrative data from Germany. Our analysis sample contains 23,679 mothers and firms, and we observe the entire employment history of all workers at those firms at the daily level. We first show that firms hire more replacements when mothers have few internal substitutes, but mothers are on average not fully replaced. Difficult to replace mothers also take substantially shorter parental leave, but this pattern is erased by a paid parental leave expansion, potentially distorting the coordination between employers and mothers. More generous paid parental leave

for medium- and high-earning mothers leads to a short-term employment reduction in firms with no long-run impacts. In a final step, we investigate whether firms adjust their hiring composition after the reform as longer birth-related absences are costly to them. We find that firms respond by hiring fewer young women, but only when they would be more difficult to replace internally. Under specific conditions parental leave regulations may accordingly have detrimental effects on (potential) mothers beyond direct labour market effects.

Chapter 4 aims to improve our understanding of enrolment gaps by parental background of children under three in the (universal) German day care system. Children from migrant or lower educated parents have been shown to benefit most in their development from attending, but these groups are underrepresented. From a policy perspective, it is crucial to understand the reasons for such gaps in order for the system to fulfil its promise of being truly universal. We use a comprehensive survey of parents with children aged 0-2 with more than 60,000 observations over five survey years. The data record both demand and usage of day care and also elicit detailed reasons for not attending. Using an Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973), we show that differences in demand play a role, but they explain only a small share of the gap, especially for migrant parents. We then analyse the role of access barriers using quasi-experimental approaches. Reducing shortages substantially decreases the gap by parental education but has no effect on the gap by migrant background. Following, we use the synthetic control method (Abadie et al., 2010) to consider the effects of a strong fee reduction in one federal state (Hamburg). As for shortages, this only reduces the gap by educational background. In sum, the gap by migrant background is much less responsive to reducing shortages or fees, and programmes to boost demand would similarly not be a silver bullet. Looking at reasons for not attending, we find some evidence for quality concerns among migrant parents. We can also not rule out that discrimination plays a role for the migrant gap.

In **Chapter 5** we estimate effects of the usage of day care on parenting activities. Across the OECD, children spend more time in day care than ever before, but at the same time parents are spending more time with their children (Dotti Sani and Treas, 2016). Understanding how parents respond to their children attending day care is important to better understand the mechanisms behind child development effects of day care. In the analysis we use time-diary and survey data from Germany. While applying a correlational analysis, we pay careful attention to selection bias and provide bounds for our estimates. We find that while parents' total time spend with their child is substantially lower, parenting activities are only reduced by a few minutes per day. During non-centre hours, parenting activities are not affected. Looking at non-

parenting activities, we see that day care is partly used to take up paid work, but also to ease time constraints. A reduction of leisure and sleep during non-centre hours while parenting activities are held constant, suggests an increase in parental motivation to maintain meaningful interactions with their child when it is attending day care.

Chapter 6 uses novel time-use data from the GDR and reunified Germany to examine the relationship between gender inequality in market work and housework. The GDR (and later East Germany) is commonly portrayed as being more gender equal than the FRG (West Germany) in both regards. Gender inequality in market work is often seen to be inherently linked to inequality in domestic work as well (Becker, 1981). In this paper we test whether this link holds beyond a pure mechanical effect. We first confirm that women in the GDR (and East Germany) indeed contribute a higher (within-couple) share of market work and a lower share of housework. However, when we decompose gender housework gaps between East and West Germany by looking at different household types and controlling for market work, we find that these are almost identical. A Gelbach (2016) decomposition reveals that more than 80% of East-West differences can be explained by women's time spend in market work. Policies in the GDR promoted female labour supply effectively, but we find no evidence that this had spill-over effects on norms regarding housework, as these are very similar in East and West Germany after reunification once individual time constraints are taken into account.

The final **Chapter 7** of this dissertation also exploits the setting of Germany's division and reunification to examine how culture impacts within-couple gender inequality. I start off by showing that the female income share in East and West German couples without children is almost evenly distributed with small East-West-differences. Using a household panel (SOEP), I estimate child penalties for couples socialised in the GDR or FRG, but who have children in reunified Germany under the same policy environment. The long-run child penalty on the female income share is 11.4 percentage points larger in West German couples (74% larger). Additionally, the arrival of children also increases the share of housework and child care performed by women more strongly in West German couples. To shed light on gender inequality in the GDR, and to compare it to East and West Germany later, I use the same time-use data from the GDR as in **Chapter 6** (these are the first two economic studies using this data) and compare the distribution of market and domestic work of couples with children to childless couples of a similar age range. In a final step of the analysis, I use another household panel (pairfam) to examine differences in attitudes. Attitudes towards maternal employment are more egalitarian among East Germans, but attitudes become more traditional for both East and West Germans when individuals have children. Exposure to different

gender norms in the GDR and FRG are shown to have strong impacts on child penalties and attitudes and are thus closely linked to overall gender inequality.

1.3 Contributions

This dissertation makes several important contributions to the literature, both as individual papers and as a whole. Contributions are discussed in greater detail in the chapters themselves, and this section serves as an overview of the main contributions. I will discuss both content- and data-related contributions.⁶ The latter may not be contributions to the literature in the classical sense, but the usage of numerous different data sets and a discussion of the advantages of those this may still provide to be useful for fellow researchers.

Chapters 2 and 3 consider the effects of different aspects of parental leave policies. Rather than looking directly at labour supply effects of such legislations, the chapters analyse how firms respond to them. Beside some recent work (Brenøe et al., 2020; Gallen, 2019; Ginja et al., 2020b), the literature on firms and parental leave is in its infancy. Potential negative effect on firms are often brought forward as a main reason against a federal parental leave policy in the US, or against more generous regimes in Europe. **Chapter 2** shows that it may be harmful for young women when firms face direct costs of motherhood. **Chapter 3** demonstrates that firm characteristics matter for the length of leave and that mothers are commonly not fully replaced by firms. Parental leave absences and more generous leave appears costly in the short-run when firms cannot easily substitute the mother internally.

The chapters also contribute to the literature on statistical discrimination against potential mothers. While direct negative effects of motherhood have been well-documented (e.g., Kleven et al., 2019a,b), young women may additionally also be negatively affected through statistical discrimination. While the concept is theoretically intuitive, most evidence on statistical discrimination stems from audit studies (Bertrand and Duflo, 2017). These chapters contribute to the scarce quasi-experimental evidence (see, e.g., Fernández-Kranz and Rodríguez-Planas, 2021) on statistical discrimination and also come with an important finding for policy; if parental leave is costly for firms, mothers and young women more broadly may be negatively affected by well-meaning policies (Blau and Kahn, 2013).

Chapters 2 and 3 use administrative linked employer-employee data from the German Institute for Employment Research (IAB). Administrative data are being increas-

⁶I will not discuss *methodological* aspects in this introduction. The chapters use a mix of descriptive and various quasi-experimental approaches (e.g., difference-in-differences, trend-break model, event-study design, Oaxaca-Blinder and Gelbach-decomposition, synthetic control method, Oster-method). I do not consider them to be direct contributions but as tools to analyse relevant research questions.

ingly used in economics and these chapters illustrate why as they allow to tackle new research questions. Analyses focusing on firms can only be insufficiently conducted using survey data.⁷ The linked administrative data allow to compute firm-level outcomes and we also observe the entire labour market history of all workers of the firms. The German data has some specific advantages, of which three were particularly important for this work; the data allow to identify single locations of multi-site firms, detailed occupational codes exist at the individual level and employment spells are at the daily level. Through this I can identify specific workgroups, which are shown in **Chapter 3** to be relevant determinants of parental leave behaviour. The fine-grained daily level allows to carefully tease out when firms conduct replacement hiring and the evolution of effects of leave-related absences for both mothers and firms can be shown in great detail.⁸ If the monthly patterns shown in, e.g., Figures 3.2 and 3.7 were aggregated at an annual level, a lot of information would be lost and the understanding of how parental leave impacts firms would be advanced to a lesser degree.

The following **Chapters 4 and 5** concern day care. Enrolment gaps for children under three by family background persist, although the system is intended to be universal. **Chapter 4** contributes to the literature on educational inequalities by family background (Björklund and Salvanes, 2011). While enrolment gaps in day care have been well documented (e.g., Drange and Telle, 2020; Magnuson and Waldfogel, 2016; Schober and Spiess, 2013; Stahl et al., 2018), the understanding of their determinants is limited. As numerous studies have shown, disadvantaged children benefit most from attending day care. Understanding the causes and designing policies to close those gaps is a crucial element to help children from all backgrounds to fulfil their potential and preventing early ingrained inequalities.

The chapter uses a representative survey from the German Youth Institute (DJI). The unusually large data for such a narrowly-defined population group (more than 12,000 parents of children aged 0-2 years are covered annually) allows to, e.g., illustrate at the monthly age when enrolment gaps appear for children (Figure 4.1), and enables a calculation of shortages at the local county-level. Without such a large, targeted survey, these types of analyses would not have been possible and advances in understanding day care enrolment gaps would have been diminished.

In **Chapter 5** we contribute to the sparse literature analysing how day care impacts on parenting activities. Development effects of day care on children are mostly linked

⁷Individual-level surveys lack sufficient firm information, firm surveys, in contrast, do not contain enough information on individual workers.

⁸In contrast, administrative data from other countries (e.g., Sweden or Denmark) are often only available at the *annual* level. Analyses in our work are conducted at the monthly level for computational reasons, but data availability enables a finer level as well.

directly to day care centres. One mechanism that is commonly neglected is that usage of day care may also directly affect the home environment (Kuger et al., 2019). We add to existing work by using detailed time-use diary data looking at detailed activities, providing a conceptual framework for centre hours and non-centre hours, and—while the chapter is correlational—we use the Oster-method (Oster, 2019) to provide bounds for our estimates. Through considering both the extensive (usage vs. no usage) and intensive (half- vs. full-day) margin we provide mechanisms that help to explain differential development effects.

Time-use diary data are particularly valuable to address this research question. Survey data with information on time use commonly only contain broadly categorised activities in hours at a weekly or daily level. In the time-use diary data we observe activities in three-digit classifications. Distinguishing between parenting activities and the child being merely present is crucial to assess how day care impacts parent-child interactions, and one can also consider specific parenting activities (e.g., reading or playing). An equally important benefit of the diary data is the inclusion of information when exactly activities are conducted (Figure 5.2 illustrates this at the aggregated hourly level). To assess how parents respond in their activities when day care is used, it is paramount to distinguish between direct and indirect effects (i.e., centre hours and non-centre hours), to understand which activities parents are trading off. With survey data only total effects could be analysed, masking the role of direct and indirect effects.

In **Chapter 6** we examine the relation between gender inequality in market and household work with time-use data from the GDR and reunified Germany. We contribute to the growing literature on gender-related differences in East and West Germany due to exposure to different policy regimes (see, e.g., Bauernschuster and Rainer, 2012; Beblo and Görge, 2018; Campa and Serafinelli, 2019; Lippmann et al., 2020). While East Germany is rightly commonly portrayed to be more gender egalitarian, we show that this only relates directly to labour market outcomes. Domestic work is on average also more equally distributed, but only due to a purely mechanical effect. Once controlling for own working time, relative contributions to domestic work are very similar in East and West Germany. Gender norms regarding market work and housework are accordingly only weakly linked (Grunow et al., 2018) and we also show that neither partner reacts strongly to labour supply of the other (Knowles, 2013). Both findings have implications on models of household bargaining.

For this research project, I and my co-authors obtained the raw files of the GDR time-use data from the German Federal Archives and after extensive data editing produced a data set of a similar structure like other international time-use studies. The GDR time-use data has only been used in the early 1990s for descriptive depictions (Priller,

1993). We are the first economic study to use this data and by making it available to the research community, future work will benefit from our efforts. The data allow for insights into gender inequality in the GDR, which is particularly valuable as other survey evidence or administrative data from the GDR are not available for researchers. In the editing process we ensured that the data is in a similar format to the time-use study of reunified Germany, through which we can make direct comparisons. The chapter also uses data from the Multinational Time Use Study (MTUS, Gershuny et al., 2020) to assess the relation between the female share of housework and paid work in a cross-country comparison. Many of the countries included in the MTUS are conducted on a household level, which enables analyses at the couple-level.

The final **Chapter 7** estimates child penalties to analyse how the arrival of children impacts within-couple gender inequality in East and West Germany. East-West-differences in gender inequality are shown to stem to a large degree from different responses to having children, an aspect that has often only been considered indirectly in this literature. At the same time, important recent contributions from different settings (Córtes and Pan, 2020; Kleven et al., 2019b, 2020) have shown that a large share of remaining gender inequality is due to children, making it important to understand determinants of child penalties. The German setting allows to better understand the role played by social norms. Related work has estimated child penalties for East and West German mothers (Boelmann et al., 2021; Collischon et al., 2020), and I add to this in several dimensions. This chapter takes a holistic view on gender inequality through considering labour market effects, but also domestic work and attitudes. Additionally, time-use data from the GDR and reunified Germany allows to compare gender inequalities in time use in a consistent way, which has to date not been possible. As this paper takes a couple-perspective, it also relates directly to other work looking at gender inequality under this lens (Angelov et al., 2016; Bertrand et al., 2015; Lippmann et al., 2020).

This chapter uses a number of different data sets to examine child-related gender inequality among East and West Germans; two household panel surveys (SOEP, pairfam) and the same time-use data from the GDR and reunified Germany that was used in the previous chapter. Using administrative data to estimate child penalties has the advantage of containing the entire labour market history of workers and large sample sizes, but an advantage of survey data in this context is that it also allows to consider domestic work as an important element of gender inequality. Additionally, attitudes can be analysed in the same empirical framework. Overall, analyses with administrative and survey data should be seen as being complementary and contributions on this topic with administrative data (Boelmann et al., 2021) and this chapter illustrate the

benefit of both.⁹ Lastly, combining the time-use data sets allows to compare gender inequality related to children in the GDR and (reunified) East and West Germany.

⁹A downside specific to German administrative data is that couples cannot be identified reliably, due to which the analyses conducted in this chapter on labour market outcomes could not be carried out.

CHAPTER 2

Punishing Potential Mothers? Evidence for Statistical Employer Discrimination From a Natural Experiment¹

2.1 Introduction

Theory predicts that employers may discriminate statistically and pay female employees of childbearing age lower wages than their male counterparts, *ceteris paribus*, if motherhood imposes costs on employers. This discriminatory behaviour will be more pronounced if firms face direct costs of motherhood as per law. We test this prediction using a natural experiment created by a 2006 reform of German maternity protection. Before the reform, large firms had to pay mothers' wage continuation around childbirth; this made their expected costs arising from maternity protection wage payments a direct function of their employees' probability to become mothers, i.e. effectively depending on gender and age of their workforce. The reform regulated that all wage continuation to mothers comes centrally from the social security system, and firms' contributions are merely a function of the number of employees and their wages, regardless of gender and age.

A driving mechanism behind the gender wage gap is the fact that, on average, mothers take substantially more time off work after childbirth than fathers. These career interruptions result in lower labour market experience and lower wages (Gangl and Ziefle, 2009; Goldin, 2014; Adda et al., 2017). However, a substantial pay gap exists even when controlling for observables such as labour market experience or tenure and an—admittedly substantially smaller—gap also exists pre-birth. One reason for this

¹This chapter is joint work with Robin Jessen (RWI and Freie Universität Berlin) and Jochen Kluge (Humboldt University Berlin, KfW Development Bank and IZA). Thanks to David Card, Mathias Huebener, seminar participants at UC Berkeley and DIW Berlin, and conference participants at EALE 2018 Lyon and COMPIE 2018 Berlin for helpful feedback. We also thank two anonymous referees and the editor of this issue of *Labour Economics*, Éric Maurin, for valuable comments and suggestions.

might be that maternity is not only punished *ex-post*, but also *ex-ante* via statistical discrimination of women of childbearing age (i.e. potential mothers). Employers in general bear some of the costs of motherhood: they need to find a replacement during leave, the accumulation of firm-specific human capital is disrupted, and existing skills of their employees deteriorate during leave. Also, mothers may not return to their job (full-time) afterwards. This can make profit-maximising companies reluctant to promote them or hire them in the first place. Under standard assumptions, disincentives to hiring potential mothers lead to adjustments along two margins: fewer potential mothers are hired, and those that are hired receive lower wages than men, *ceteris paribus*. In theory this effect is stronger if companies have to pay wage continuation directly. Using a natural experiment, we estimate to what extent this effect can be counter-acted if the state pays wage continuation centrally through the social security system.

Before January 1, 2006, large firms in Germany (more than 30 employees) were obliged to pay for the generous maternity protection—14 weeks of 100 per cent wage continuation around the date of delivery—of their female employees. From 2006 onward, each firm contributed to maternity protection through a contribution to the countrywide health insurance system, which is paid for every worker in the firm, irrespective of whether the worker is male or female (i.e. firms with an all-male workforce contribute the same as firms with an all-female workforce). This had been the regulation for small firms (≤ 30 workers) already before the reform. In its ruling declaring the previous regime as unconstitutional, the German Federal Constitutional Court stated it was unable to say with certainty "how large the probability is that due to this burden employers who have to pay maternity protection wage continuation decided not to employ women" (Bundesverfassungsgericht, 2003a, Section 120). However, not only hiring decisions but promotions and, more generally, wage setting may be affected. We use comprehensive data—from annual linked employer-employee administrative records—covering over 1 million workers in more than 10,000 firms to estimate the wage effects of the reform.

Gender wage and employment differences have received extensive coverage in the literature. Despite signs of a narrowing over time these differences have persisted in all Western economies, see, e.g, Blau and Kahn (2003), Weichselbaumer and Winter-Ebmer (2005), Olivetti and Petrongolo (2008) and OECD (2019). Card et al. (2015) look at firm-specific pay premiums as a source of (gender) wage inequality and point out that if firms have some control over the wages offered, relative wages of women will be influenced by both a potential sorting of women into higher or lower paying firms and on their relative bargaining power. In Germany, the share of females is higher in

small firms, which on average pay lower wages. Heinze and Wolf (2010) confirm some selection of women into lower paying firms more generally, i.e. for firms of all sizes.

To our knowledge, no previous study links the gender wage gap to *statistical* discrimination of *potential* mothers. So far, the literature has focused on *actual* mothers.² However, not only *being* a mother but already the *possibility of becoming* a mother can have adverse effects in the labour market. This links the topic to the literature on statistical discrimination, which predicts wage differences between groups even in the absence of taste-based discrimination as in Becker (1957). The basic idea of statistical discrimination—originally developed by Phelps (1972)—is that profit-maximising employers have imperfect information about the productivity of potential employees and rationally use group statistics as proxies for unobserved exogenous characteristics (see Fang and Moro, 2011).³ In the case of statistical discrimination against potential mothers, firms do not know whether a specific woman will become a mother, but use the information that a share of women of childbearing age will have children (about 80% in Germany), which imposes additional expected costs on firms.⁴ Therefore the firm is expected to be willing to pay women of childbearing age a lower wage than men of same age. This is a case of statistical discrimination, because even women who will not have a child are paid a lower wage than men due to the use of group statistics for gender and age to determine expected costs due to childbearing.

While in Germany anti-discrimination laws regulate that employers are not allowed to ask (potential) employees about a current or planned pregnancy, it is an evident possibility that a woman of childbearing age will become a mother: the annual average probability to give birth for working women in the age bracket 24-35 is 5.5%, peaking at an average of 7% p.a. for women aged 30-32. Adverse labour market prospects for women can then arise without any taste-based discrimination, and, as Phelps (1972, p. 661) notes, "[d]iscrimination is no less damaging to its victims for being statistical."

²A recent study by Biewen and Seifert (2018) is an exception: while not focusing on the gender wage gap, it quantifies the association of the probability of parenthood on career transitions for men and women in Germany, and finds a negative relationship between the contemporaneous probability to have a child and horizontal career transitions for women. Horizontal career transitions are defined as job changes in which the number of subordinates does not change by more than two. These transitions might thus still be associated with substantial wage increases.

³Arrow (1973) develops a similar model in which, in contrast, average group differences in characteristics are endogenous.

⁴Specifically, before the reform the costs per hour worked of a mother were higher for large firms, because the amount of hours worked in the year a woman had a child were reduced due to maternity protection, but firms' expenses (wage and wage continuation) were not reduced proportionally. In contrast, small firms did not have to pay wage continuation during maternity protection and therefore the expenses per hour worked are the same regardless of whether a woman has a child. See also Section 2.3.

In the empirical analysis, we implement a *difference-in-differences (DID) model* to capture the post-reform effect, as well as a *trend-break (TB) model* that distinguishes between general convergence (or divergence) in female and male wages at large firms over time and the divergence from this trend post-reform. The treatment effect would imply, and measure, the corresponding degree of statistical employer discrimination pre-2006.

Our results confirm the theoretical prediction: the DID estimates show a statistically significant increase in female wages relative to male wages of 1.1 per cent in the post-reform period. The magnitude of the effect is in line with a back-of-the-envelope calculation, which puts the expected effect at around 0.8 per cent. The trend-break model carves out this pattern in more detail and shows that the general time trend in the convergence of the gender wage gap in large firms is not statistically significantly different from zero and the post-2006 trend estimates a total reform effect of 2.5 per cent over five years. Hence, the estimated effect confirms the hypothesis that the statistically discriminatory behaviour of large firms pre-reform reflects the costs for the firm arising from the regulation. Overall, the findings imply that policy makers should try to identify factors that could result in statistical discrimination against potential mothers. If the aim is to prevent negative labour market effects *ex-ante* and *ex-post* for (potential) mothers, it is worthwhile for the public to compensate firms for costs that occur asymmetrically to mothers only.

2.2 Institutional background

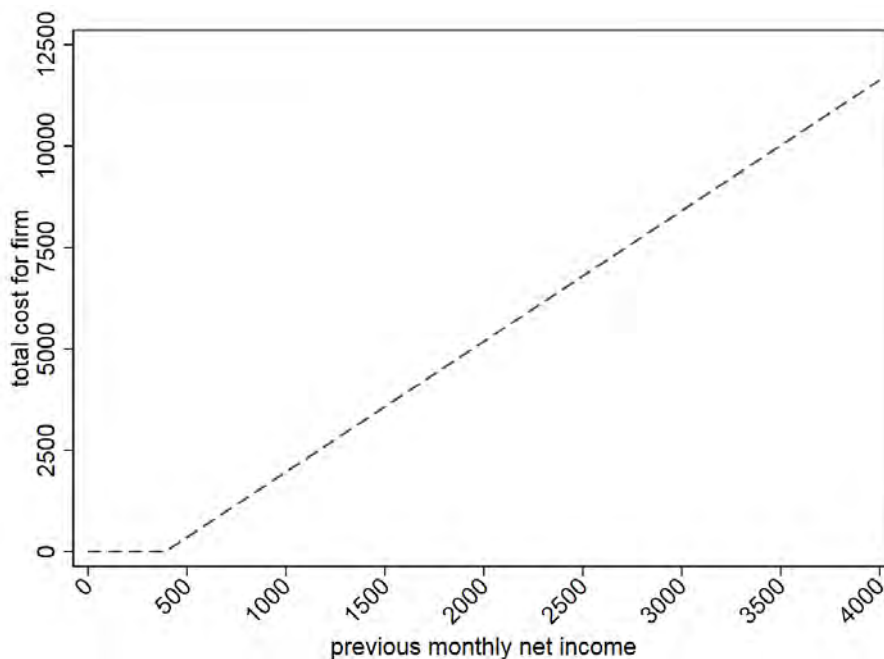
Mothers in Germany are entitled to paid maternity protection, which lasts from six weeks before the predicted date of childbirth until eight weeks after. During those 14 weeks expecting mothers are entitled to 100 per cent wage replacement. This paid maternity protection regulation is the focus of our analysis. In addition, several other regulations are in place that protect mother and child and the mother's employment relationship: these include, for instance, dismissal protection starting with the onset of the pregnancy, and job-protected maternity leave for up to three years post-childbirth with the option to return to the same job.

Before 2006, employers in Germany with more than 30 full-time-equivalent employees had to pay a substantial share of the wage replacements of mothers during maternity protection themselves:⁵ women who were insured by a statutory health insurance com-

⁵The regulation originally set the threshold at 20 and granted the statutory health insurers the flexibility to increase that threshold from 20 to up to 30. We contacted the different regional entities of the largest statutory health insurer, AOK, which were responsible for executing the maternity protection payments. The vast majority of them set the limit to 30, therefore we use this effective cut-off in our analysis to ensure that all firms are correctly assigned. The relevant measure and threshold of

pany received a fixed amount of 13 Euro per calendar day (roughly 400 Euro monthly) and additionally, if they were employed previously, the difference to their previous net earnings from their employer. In 2001, firms paid a total of about 1.48 billion Euro for maternity protection wage continuation (Bundesverfassungsgericht, 2003a).⁶

Figure 2.1: Firms' costs per maternity protection period



Notes: Costs of maternity protection imposed on firms for different levels of employee's pre-birth monthly net income in Euro.

Figure 2.1 displays the costs firms faced for the 14 weeks of maternity protection as a function of previous net earnings. From a monthly net income of about 400 Euro the costs are a linearly increasing function of the previous net wage. Note that the figure displays firms' direct, effective total costs for one female worker in maternity

full-time equivalent employees (FTE) is determined by the number of employees weighted by hours worked. A person working less than 10 hours counted 0.25 FTE, a person working 10 to 20 hours 0.5 FTE, 20-30 hours 0.75 FTE and a person working more than 30 hours counted as 1 FTE. In our data (see Section 2.4) a part-time variable indicates whether a person has worked more or less than 18 hours per week, and the FTE categories are approximated. We use the German Socio-Economic Panel (SOEP, Goebel et al., 2019) to impute the respective shares of workers falling into the working hour groups. We use information from 2003, in line with our specification of small and large firms (see below), and implement the imputation separately by gender.

⁶Note that individuals with private health insurance were not covered by this regulation and also not affected by the change in law. However, since only about 10 per cent of Germans are privately insured, and since our data do not contain civil servants (see Section 2.4), of which many have a private insurance, this is not a major issue. Concerning our estimates, we still identify lower bounds in absolute terms for individuals covered by public health insurance, as we cannot determine the type of insurance an individual had.

protection: the fixed share of 13 Euro per calendar day (a total of about 1,400 Euro for the 14 weeks) covered by the statutory health insurance is deducted, and the cost curve begins its upward sloping part at monthly earnings of about 400 Euro. As this 13 Euro daily contribution paid by insurers had remained unchanged since 1968, whereas wages had risen substantially, the share paid by large firms had increased strongly over time.

For small firms, in contrast, statutory health insurance companies paid the entire wage continuation. These firms had to pay a social security contribution per employee to compensate for the costs, a pay-as-you-go system called *Umlage U2—Mutterschaft* (Contribution U2—Motherhood). Small firms had to contribute to *U2* per worker regardless of gender and age, a set-up explicitly designed to prevent adverse employment effects for women of childbearing age. In 2003, around 90 per cent of all firms were covered by *U2*, but this included only one-third of female and one-quarter of male employees.

Following a constitutional complaint of a large firm, on 18 November 2003 the German Federal Constitutional Court ruled the legislation as unconstitutional. The judges argued that, since it created a disincentive for large firms to hire women, it violated the constitutional principles of equal opportunities for men and women and the right to choose the workplace freely (Bundesverfassungsgericht, 2003b). The court demanded that a new regulation should be conceived in line with the principles of the constitution. Theoretically, anticipation effects might have played a role from the moment of the ruling onward, although the court gave the government flexibility concerning the specific design of the legislation. Eventually on January 1, 2006, a new law came into effect regulating that firms of all sizes have to take part in the pay-as-you-go system *U2*. That is, large firms are treated in the same way as small firms; they have to pay the social security contribution for maternity protection payments—a pure function of the number of employees and their wages, irrespective of gender and age composition—and the statutory health insurance companies reimburse firms for the wage replacements.

2.3 Mechanisms and expected effects

This section first explains how the pre-reform regime influenced large firms' expected costs per week of work of potential mothers. Second, it describes the mechanism through which the 2006 reform is expected to impact on wages of female and male workers at large firms.

When calculating the costs of employing a potential mother, employers take into account the firm's expected total costs of motherhood. These include costs that were not affected by the 2006 reform, including, e.g., finding a replacement for the mother

while she is on leave, or reductions in output during this period. The reform did affect whether companies had to contribute to the 14 weeks of wage continuation during maternity protection. When firms contribute, this increases their costs per week worked of mothers. If male and female workers are perfect substitutes, profit maximising firms are expected to pay potential mothers such that the expected costs per week worked are the same as those for men of the same age, *ceteris paribus*. The magnitude of total wage continuation $Cont$ paid by large firms before the reform is given by Equation (2.1):

$$Cont = \begin{cases} 0 & \text{if } X(W) \leq 13 \times 365 \\ X(W) \times 14/52 - 13 \times 14 \times 7 & \text{if } X(W) > 13 \times 365, \end{cases} \quad (2.1)$$

where W denotes the annual gross wage and $X()$ is a function that converts gross wages to net wages. 13 Euro per day of the 14 weeks of maternity protection were paid by insurers.

Denote the weekly gross wage by w . The *expected* effective costs per week of work including employer social security contributions of a potential mother are given by

$$c_{exp} = p \left(\frac{Cont}{weeks} + w + SSC \right) + (1 - p)(w + SSC), \quad (2.2)$$

where SSC denotes employer social security contributions per week, about 20 per cent of the gross wage, p is the probability that a potential mother will give birth during her current contract, 5.5% per year in our sample,⁷ and $weeks$ are the weeks effectively worked under the current contract. For instance, if the mother has a one year contract and returns to work immediately after maternity protection, this figure is 38.⁸ Note that the effective costs per week of work for a mother under the old regime increase substantially if she takes a long parental leave and thus works fewer weeks under her contract. In the post-reform regime, where $Cont = 0$, the effective costs per week are independent from the length of parental leave. The first term of Equation (2.2) gives the effective costs per week if the woman gives birth multiplied by the probability of this event. The second term gives the probability of not giving birth multiplied by the weekly wage including social security contributions.

As a benchmark for the magnitude of the treatment effect, consider a woman whose monthly gross earnings equal the average in our sample, 2,440 Euro: given the 2006 tax regime and single filing her labour income net of taxes and social security contributions

⁷We implement a method developed by Müller and Strauch (2017) to identify births in German social security data. The annual births probabilities are in line with those calculated by Raute (2019), but slightly higher as we restrict the upper age limit to 35 in our calculations.

⁸Adjusting 52 weeks per year for vacations would increase the terms $Cont/weeks$, w , and SSC by the same proportion. Thus it would not change the ratio of the expected effective costs for potential mothers pre-reform and post-reform.

is 1,547 Euro.⁹ Using Equation (2.2) and assuming a probability of giving birth of 5.5% and that she would return to work directly after maternity protection, the expected costs per week of work are 681.08 Euro. In the post-reform regime, where $Cont$ is zero and the equation simplifies to $w + SSC$, the costs per week of work would be 675.69 Euro, about 0.8 per cent lower. In practice, women on average take considerable time off after motherhood, therefore this figure is a lower bound for the change in expected costs. Thus, if the gains of the reform are entirely reaped by female workers through higher wages, one would expect wages of potential mothers to increase by at least 0.8 per cent as a result of the reform.

In practice, the adjustment processes may affect both wages and the share of employed women. The reform decreased the expected costs of employing female workers. At constant wages, this leads to firms hiring more women. If the elasticity of substitution of male and female workers equals unity, i.e. they are neither substitutes nor complements, and abstracting from other inputs in the production function, the ratio of expenditure on female and male workers is unchanged and thus the number of employed males remains constant. If instead men and women are substitutes, the number of employed males decreases. If they are complements, the number of both males and females will increase and the sign of the impact on the ratio of male and female workers is ambiguous. The effects on wages depend on the bargaining power of women. If female workers have at least some bargaining power, their wages will increase if their labour costs decrease. Then if female and male workers are substitutes, female wages will increase relative to male wages.

2.4 Data and empirical approach

The empirical analysis uses linked employer-employee data taken from social security records from the German Institute for Employment Research (IAB). The data set is constructed by first selecting a sample of establishments that took part in the IAB's annual establishment survey and then selecting individuals who have worked at any of these establishments for at least one day between 2002 and 2012. For those individuals we observe the full employment biography from 1993 to 2014 (Heining et al., 2016). As the employment data are based on social security records, civil servants, self-employed and individuals in higher education are not covered.

The individual spell data include detailed information on workers' employment history and gross daily wages or benefits and contain a limited set of socio-demographic characteristics such as gender, age and other variables with less complete coverage. By

⁹This figure can be obtained using the tax calculator provided by the Federal Ministry of Finance and applying social security contributions of 0.2 of gross earnings.

law, employers are obliged to report the beginning and end of each employment relationship along with an annual report at the end of each year. The reported daily wage comprises all gross earnings including premiums and allowances. As is common for social security data, the wage is censored from above at the upper earnings limit in the statutory pension fund. The threshold is adjusted in most years and differs between East and West Germany, e.g. in 2010 the threshold was 66,000 Euro in West Germany and 55,800 Euro in East Germany. The data also contain information on full-time, part-time and marginal employment, but do not contain more detailed information on hours of work. Hence one cannot conclude unambiguously whether a change in daily earnings is due to a change in the hourly wage or in hours worked.

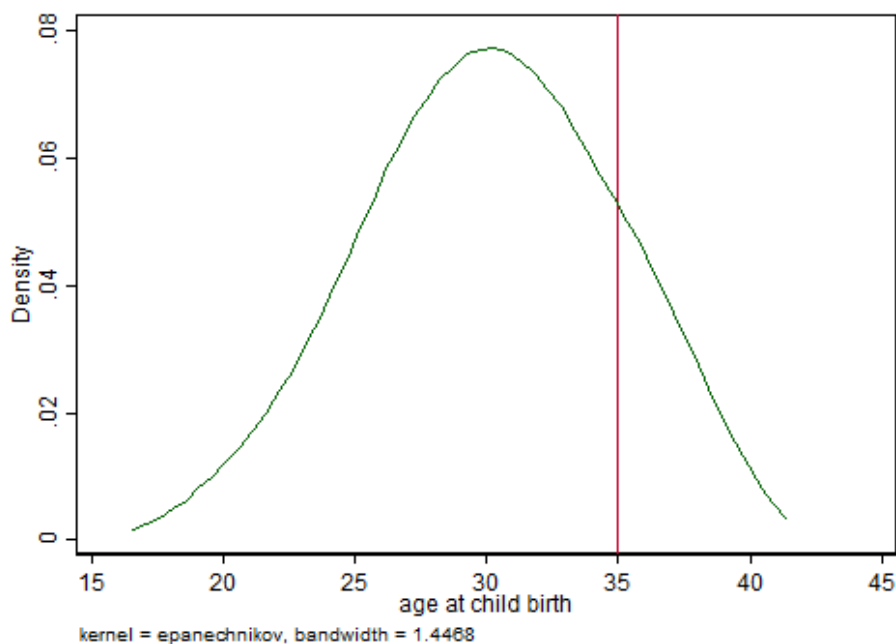
At the establishment level, we observe the sector, the geographical location as well as the founding and—if applicable—closing year. In total we observe more than 10,000 establishments in each year, which have been linked to individuals through a unique establishment ID, corresponding to more than 1 million individual observations per year.

We exclude person-level observations with daily wages below 1 Euro, which we assume are either dormant employment relationships or spells attributable to measurement error. We keep both regularly and marginally employed individuals, because the definition of marginal employment changed in 2003, transforming some regular employment relationships into marginal ones. Excluding marginally employed individuals would lead to a drop in the observed average wage in 2003. We include only individual spells that cover June 30 of a given year, as this is the point in time when the surveys at the establishments are conducted. Finally we exclude firms with nine employees or fewer when briefly contrasting small and large firms, since these very small firms arguably are too different from large ones to constitute an adequate comparison. All analyses restrict the sample to individuals up to 35 years of age. This age threshold is chosen based on the childbirths identified in our sample: the age distribution of mothers at birth is plotted in Figure 2.2 and indicates that after the age of 35 only a small share of women is likely to give birth in the future.

We conduct the analysis of (full-time) wages at the individual level.¹⁰ We use the cut-off of 30 full-time equivalent employees defined by the maternity protection legislation before 2006 to ensure that firms were affected by the change in law. Assignment to the group of large firms is based on firm size in the year 2003, prior to the ruling of the constitutional court. Note that this implies that individuals can move between large

¹⁰Note that this is equivalent to conducting an analysis of the gender wage gap at the firm level and weighting annual firm observations by the number of employees used to calculate average wages at the firm level.

Figure 2.2: Density plot of mothers' age at childbirth



Notes: Authors' calculations based on the administrative data described in Section 2.4; births are identified using a method developed by Müller and Strauch (2017). Only childbirths for mothers who were employed subject to social security contributions (pre-birth) are covered. Pooled over the years 2001 to 2010.

and small firms. Therefore, effects on wages could reflect a composition effect to some extent.

In principle, one alternative would have been to assign workers to whether they are in a small or large firm (i.e. not part or part of the analysis sample) depending on the size of their firm in 2003. However, this leads to a substantial problem; one could then only include individuals who were employed in 2003, which would lead to an ageing sample by construction and therefore one could not analyze the effect on women of childbearing age in later years. While the research design for testing the hypothesis on statistical discrimination is thus based on the assignment to the analysis sample conditional on the firm size, we make use of the advantages of merged worker-establishment data by combining individual and firm information (and fixed effects) to analyse individual wages conditional on the size of the establishment.

Table 2.1 displays summary statistics of small and large firms. We contrast firms of different sizes to emphasise the relevance of looking at large firms. As these pay higher wages, improving employment prospects for women at larger firms can be an effective way to reduce the overall gender wage gap in the economy. We focus on the years 2001 to 2010 in the analysis in order to take into account a sufficiently long enough

Table 2.1: Summary statistics: firms by size and sector

	All firms		Manufacturing		Service		Public sector		Other sectors	
	10-30	> 30	10-30	> 30	10-30	> 30	10-30	> 30	10-30	> 30
Number of employees										
Monthly full-time wage, women	1859 (708)	2773 (675)	1714 (640)	2947 (670)	2224 (799)	2790 (722)	2030 (644)	2397 (474)	1470 (602)	2496 (627)
Monthly full-time wage, men	2105 (731)	2963 (684)	2045 (592)	3110 (630)	2537 (1026)	3085 (804)	2226 (774)	2631 (626)	1755 (607)	2539 (615)
Share of female employees	0.40 (0.35)	0.35 (0.25)	0.27 (0.30)	0.22 (0.15)	0.50 (0.31)	0.48 (0.19)	0.66 (0.33)	0.66 (0.19)	0.31 (0.31)	0.23 (0.19)
Share of female new hires	0.41 (0.41)	0.39 (0.26)	0.28 (0.37)	0.27 (0.19)	0.49 (0.38)	0.51 (0.23)	0.67 (0.37)	0.66 (0.22)	0.37 (0.38)	0.26 (0.24)
Share of women working full-time	0.69 (0.37)	0.82 (0.20)	0.73 (0.37)	0.87 (0.14)	0.75 (0.35)	0.79 (0.22)	0.57 (0.37)	0.67 (0.23)	0.71 (0.37)	0.86 (0.20)
Share of men working full-time	0.87 (0.27)	0.93 (0.16)	0.92 (0.19)	0.98 (0.08)	0.83 (0.31)	0.89 (0.20)	0.73 (0.38)	0.82 (0.23)	0.87 (0.26)	0.94 (0.14)
Share women of FT employees	0.35 (0.29)	0.28 (0.23)	0.24 (0.24)	0.17 (0.13)	0.43 (0.24)	0.37 (0.17)	0.57 (0.29)	0.56 (0.20)	0.29 (0.23)	0.18 (0.16)
Number of new hires	3.5 (17.4)	214.5 (480.9)	2.2 (3.7)	293.5 (588.6)	9.6 (45.4)	122.4 (169.9)	2.7 (4.4)	94.4 (209.0)	3.8 (5.7)	46.6 (48.5)
At least one new hire (=1)	0.70 (0.46)	0.98 (0.14)	0.69 (0.46)	0.98 (0.13)	0.72 (0.45)	0.98 (0.14)	0.71 (0.46)	0.98 (0.15)	0.74 (0.44)	0.96 (0.20)
At least one woman hired (=1)	0.41 (0.49)	0.91 (0.29)	0.30 (0.46)	0.89 (0.31)	0.52 (0.50)	0.95 (0.22)	0.59 (0.49)	0.96 (0.2)	0.43 (0.50)	0.79 (0.41)
Individual-year observations	47,466	1,139,324	26,213	700,156	6,965	147,259	8,506	234,900	5,782	57,009
Firm-year observations	10,658	17,822	5,428	8,679	1,458	2,209	2,539	5,648	1,233	1,286

Notes: Table entries are means and standard deviations (in parentheses) of the variables. All variables are calculated using only employees below 35 years of age. Bold variables indicate dependent variables in the analysis. New hires refer to employees who joined the firm in the 12 months up to 30 June of a given year. Unit of observation for summary statistics is the firm, pooled over the years 2001-2010. Individual-year observations refer to full-time employees which are used in the analysis. Number of employees denotes full-time equivalent employees. Sectors are based on the 2003 edition of the German Classification of Economic Activities (WZ 2003) at the firm level. Manufacturing consists of codes D-G, service of K and L, public sector of N-Q, other sectors capture the residual sectors (agriculture, forestry, mining, hospitality industry and transportation).

period prior to the 2006 reform (and the ruling of the constitutional court); at the same time, given the mechanisms laid out in the previous section it seems plausible that the treatment effect will have materialised and the adjustment process completed within five years post-reform. The table also presents summary statistics by main sectors, in particular classifying workers and firms into manufacturing, services, the public sector and a residual of other sectors using a three-digit sector classification. The firm level summary statistics are based on our analysis sample (individuals below 35 years).

Looking at the mean monthly gross full-time wage, several pronounced patterns are apparent: first, male average monthly wages are consistently higher than female average wages, with an absolute difference of around 200 to 300 Euro, around 10 per cent in relative terms. Second, for both men and women, average wages are substantially higher in large firms than in small firms: in the full sample, average wages at large firms are about 915 Euro higher than at small firms (approx. 50 per cent) for women and about 860 Euro higher (approx. 40 per cent) for men. These differences, in turn, vary significantly by sector. In the manufacturing sector, the large-small firm differential is more than 1,000 Euro in absolute terms (for both men and women), in service occupations it is around 550 Euro, and in the public sector it is around 400 Euro.

The average share of female workers is lower at large firms (35 per cent) than at small firms (40 per cent). The main pattern of the share of female workers by firm size is maintained across the main three sectors shown in Table 2.1, albeit at different levels: females make up about one fourth of the workforce in manufacturing firms, about half of the workforce in service firms, and about two thirds in the public sector. Table 2.1 also shows that the average share of female workers among new hires is about 40 per cent in both large and small firms. Again, differences are visible by main sector: in manufacturing, the share of new hires who are female is just over one fourth, in service occupations it is about half, and in the public sector it is two thirds.

The share of female workers working full-time, unsurprisingly, is consistently lower than the share of male workers working full time. *Within* gender, i.e. looking at the share of women in the firm who work full-time, this share is substantially *higher* in large firms than in small firms: the difference amounts to 13 percentage points in the full sample (82 per cent vs. 69 per cent); it is of similar size in the manufacturing (87 vs. 73 per cent) and public sectors (67 vs. 57 per cent), but smaller in services occupations (79 vs. 75 per cent). *Across* gender, however, i.e. looking at the share of women among all full-time employees, this share is typically *lower* in large firms than in small firms.

We estimate the causal effect of discontinuing the regime of letting large firms pay for maternity protection wage continuation on wages of women up to 35 years of age at large firms using two specifications. As shown in Table 2.1, wages of men and women differ fundamentally. A DID specification allows for differences among groups, but crucially relies on the common trend assumption, i.e. in the absence of the reform the groups would have followed the same trend and the difference between them would have remained constant. By definition this assumption cannot be tested. The following specification gives the DID estimator for the outcome variable log monthly wages $\ln(w)$ for individual i in firm j at time t :

$$\ln(w_{ijt}) = \gamma_1 \text{female}_i + \gamma_2 \text{female}_i \text{post}_t + \text{year}_t + \omega_j + \epsilon_{ijt}. \quad (2.3)$$

year_t and ω_j are year and firm fixed effects, and the indicator female_i equals one for women. post_t is a binary indicator that takes on the value one from 2006 onward. The coefficient γ_2 is the DID estimator, and ϵ_{ijt} is the error term. Potential control variables at the firm level are limited in the administrative data and candidates (e.g. founding year, state, or sector) are absorbed by firm fixed effects.

For the comparison of male and female wages the common trend assumption of the DID estimator might be problematic: instead of a common trend there could be some convergence over time, even in the absence of the reform. We thus propose

an alternative estimator, which nests DID. Specifically, we assume that the rate of convergence (or divergence) in the outcome between treatment and control group would be constant in the absence of the reform. The reform effect manifests itself in a break of this trend. The *trend-break model* is specified as

$$\ln(w_{ijt}) = \gamma_1 \text{female}_i + \gamma_2 \text{female}_i \text{post}_t + \delta_1 \text{female}_i \text{trend}_t + \delta_2 \text{female}_i \text{posttrend}_t + \text{year}_t + \omega_j + v_{ijt}. \quad (2.4)$$

Here $\text{posttrend}_t = \text{year}_t - 2005$ if $t > 2005$ and zero otherwise and $\text{trend}_t = \text{year}_t - 2000$. The coefficient δ_1 thus gives the annual convergence or divergence between the control and treatment groups, and the coefficient δ_2 gives the diversion from this longer-term trend for the post-reform years. In case of a common trend in the years before the reform, δ_1 would equal zero. The DID estimator γ_2 indicates a one-time shift in the outcome variable due to the reform. δ_2 indicates the additional average *annual* effect of the reform on the treated. For instance, the effect in the first post-reform year is $\gamma_2 + \delta_2$, in the second year it is $\gamma_2 + 2 \times \delta_2$, etc. The TB specification has two advantages. First, it does not rely on the common trend assumption, and second, it may be more in line with the expectation that wages adjust gradually to new regimes.

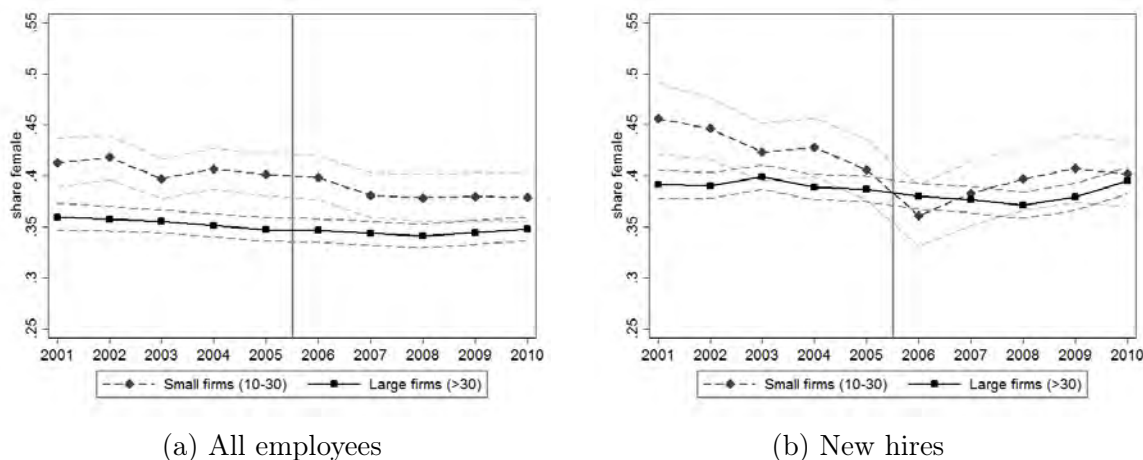
2.5 Results

This section presents the empirical results. Figure 2.3 shows the share of female employees in large and small firms for all employees and for newly hired employees. Two things are worth pointing out; first, the share of female employees in large and small firms does not appear to follow a common trend pre-reform.¹¹ Second, both the female employment share and the female share in hiring are virtually constant in large firms before and after the reform. An interesting finding in itself, the latter gives the wage analysis within large firms more legitimacy, as concerns about a compositional bias in the wage estimations are alleviated. The clearest adjustment mechanism—and, hence, evidence for statistical employer discrimination—can be observed by contrasting female and male wages *within* large firms. Moreover, this analysis is not contaminated by potential spill-overs from large to small firms.

Figure 2.4 begins with a descriptive investigation, plotting log full-time monthly wages for men and women and the respective female-male differential in large firms over the observation period. Panel (a) of the figure shows a parallel—i.e. flat—development of the respective male and female wages for all employees during the pre-reform period,

¹¹We explored an analysis of the reform effect using the share of female employees in large vs. small firms as a dependent variable, but due to the lack of a common trend this analysis was inconclusive.

Figure 2.3: Share of female employees in small and large firms, 2001-2010



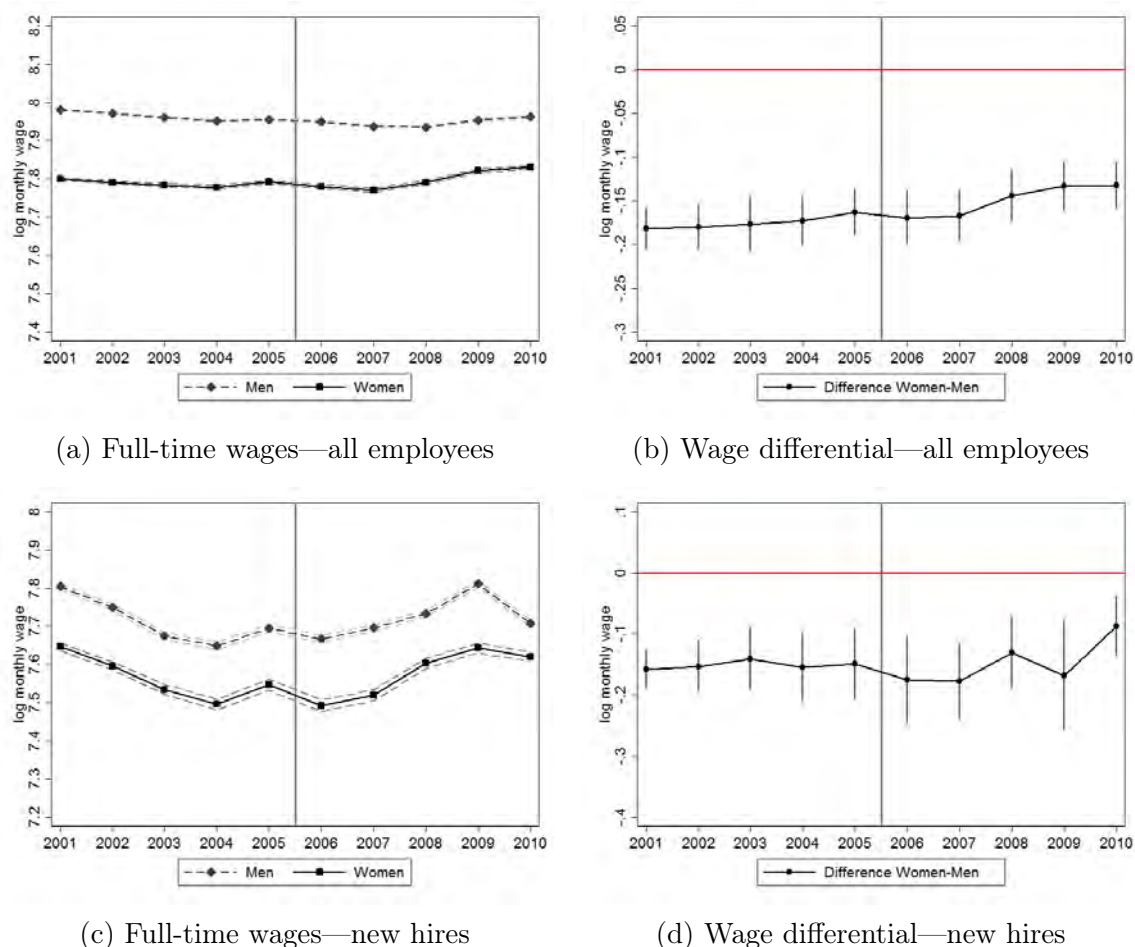
Notes: Figures display the share of female employees in small (10-30 employees) and large firms (>30). Panel (a) plots shares for all employees up to 35 years of age, Panel (b) for newly hired employees specifically. The vertical line indicates the reform. Firms are weighted by their number of full-time equivalent employees. Light grey dashed lines indicate 95 per cent confidence intervals.

and an increase in female wages vs. a continued flat curve for male wages during the post-reform years. Panel (b) illustrates this pattern for the gender wage differential, and again shows its narrowing during the post-reform years. The figure therefore gives some indicative evidence for the expected adjustment mechanism and statistical discrimination by employers.

When looking specifically at wages for new hires in the bottom panels, given the resulting smaller sample sizes, the patterns shown are somewhat noisier than the top panel. Panel (c) indicates a slightly decreasing trend in wages for both male and female new hires—a trend that would possibly be in line with a German labour market that was at the time (early to mid 2000s) characterised by the highest unemployment rate since the 1950s and by wage moderation (the same trend is also visible for female wages in small firms, which are not reported here). Perhaps more importantly for this study, however, the bottom panel of Figure 2.4 shows that during the pre-reform period also for new hires male and female wages in large firms display a parallel development and that the corresponding gender wage differential is narrowing slightly during the post-reform years.¹² The confidence intervals in Panel (d) indicate that this is less precisely estimated than for all workers, but the main pattern is still visible.

¹²One might wonder whether the overall macroeconomic trend—decreasing employment rates before 2006, increasing employment rates from 2006 onward—might affect wage growth differentially for men and women. Burda and Seele (2017) show that, while the overall downward trend in real wages between 2003 and 2011 is the same for men and women, it is slightly more pronounced for women. If anything, this would make our analysis underestimate the reform effect on wages.

Figure 2.4: Average full-time wages of men and women in large firms (> 30 employees), 2001-2010



Notes: Figures on the left display annual averages of log monthly wages for female and male full-time employees up to 35 years of age in large firms (>30 employees). Light grey dashed lines indicate 95 per cent confidence intervals. Figures on the right show the coefficient for the difference between the groups and 95% confidence intervals. Panels (a) and (b) plot wages for all employees, Panels (c) and (d) for newly hired employees. The vertical line indicates the reform.

The estimation results in Table 2.2 indicate that the adjustment process in large firms anticipated in theory has indeed taken place in practice: first, the DID estimate in column (1) shows an increase in female wages relative to male wages of 1.1 per cent, statistically significant at the 1 per cent level. The TB model in column (2) carves out the pattern more finely: the general time trend in the convergence (or divergence) of the gender wage gap in large firms is not statistically different from zero. The estimated reform effect in the TB model is composed of the coefficients γ_2 and δ_2 . That is, the reform effect is estimated to be -0.56 per cent in 2006 ($1 \times 0.76 - 1.32$),

Table 2.2: Estimates of the reform effect on log full-time wages within large firms

	Full sample		Above median wage		Below median wage	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All employees						
Post-period \times female (γ_2)	0.0113*** (0.0040)	-0.0132*** (0.0043)	0.0098* (0.0054)	-0.0142** (0.0055)	0.0143** (0.0063)	-0.0126* (0.0066)
Time trend \times female (δ_1)		0.0004 (0.0016)		-0.0013 (0.0019)		0.0039* (0.0020)
Post-2006 trend \times female (δ_2)		0.0076*** (0.0024)		0.0102*** (0.0029)		0.0026 (0.0030)
Clusters	2,007	2,007	528	528	1,477	1,477
Observations	1,139,324	1,139,324	622,968	622,968	516,301	516,301
Panel B: Newly hired employees						
Post-period \times female (γ_2)	0.0151* (0.0091)	-0.0074 (0.0152)	0.0207 (0.0156)	0.0060 (0.0281)	0.0118 (0.0099)	-0.0155 (0.0128)
Time trend \times female (δ_1)		0.0011 (0.0036)		-0.0070 (0.0067)		0.0053 (0.0040)
Post-2006 trend \times female (δ_2)		0.0056 (0.0046)		0.0167** (0.0079)		0.0002 (0.0061)
Clusters	1,975	1,975	521	521	1,452	1,452
Observations	217,202	217,202	99,952	99,952	117,235	117,235
Year FEs	Y	Y	Y	Y	Y	Y
Firm FEs	Y	Y	Y	Y	Y	Y

Notes: Table entries are the interaction terms of Equations (2.3) and (2.4). Dependent variable is log monthly wage. Sample is restricted to full-time employees up to 35 years of age. Observations denote individual by year observations. Columns (1) and (2) use the full sample, columns (3)-(6) stratify according to the median wage among all large firms in 2003. Standard errors (in parentheses) are clustered at the firm level. */**/** denote statistical significance at the 10/5/1% level, respectively.

0.2 per cent in 2007 (2 x 0.76-1.32), 0.96 per cent in 2008, 1.72 per cent in 2009, and 2.48 per cent in 2010. The latter value is the total reform effect after five years. An F-test yields that the estimated reform effect is statistically significant at the 5 per cent level from 2008 onward and insignificant before. This result indicates that indeed under the pre-2006 regulation employers showed statistically discriminating behaviour against women, depressing their wages, which was then reversed with the reform.

The bottom panel of Table 2.2 presents estimation results for the wages of new hires. The point estimates for the DID (column (1)) and the TB models (column (2)) are similar in size to those for all employees, but less precisely estimated. The DID point estimate indicates a positive reform effect of 1.5 per cent (significant at the 10 per cent level) on wages of newly hired female workers. Columns (3) through (6) of Table 2.2 investigate further these reform effects, by distinguishing large firms that pay high wages from large firms that pay lower wages: specifically, we classify firms as above and below median wage if their average wage was above or below the sample median in the year 2003 (such that this classification, again, is independent of the reform).

Table 2.3: Estimates of the reform effect on log full-time wages in large firms—by sector

	Manufacturing		Service		Public sector		Other sectors	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All employees								
Post-period \times female (γ_2)	0.0209*** (0.0048)	-0.0003 (0.0061)	0.0174** (0.0075)	0.0041 (0.0110)	0.0176** (0.0079)	-0.0034 (0.0095)	0.0113 (0.0155)	0.0116 (0.0199)
Time trend \times female (δ_1)		0.0017 (0.0024)		-0.0015 (0.0029)		0.0009 (0.0022)		0.0029 (0.0062)
Post-2006 trend \times female (δ_2)		0.0044 (0.0031)		0.0070* (0.0041)		0.0053 (0.0036)		-0.0050 (0.0062)
Clusters	1,004	1,004	259	259	599	599	145	145
Observations	700,156	700,156	147,259	147,259	234,900	234,900	57,009	57,009
Panel B: Newly hired employees								
Post-period \times female (γ_2)	0.0207 (0.0130)	0.0202 (0.0187)	0.0343* (0.0180)	0.0067 (0.0360)	0.0245 (0.0158)	-0.0103 (0.0229)	-0.0200 (0.0202)	-0.0163 (0.0462)
Time trend \times female (δ_1)		-0.0135* (0.0077)		0.0045 (0.0069)		0.0127** (0.0059)		-0.0062 (0.0121)
Post-2006 trend \times female (δ_2)		0.0233** (0.0106)		0.0017 (0.0068)		-0.0099 (0.0083)		0.0090 (0.0114)
Clusters	994	994	254	254	586	586	141	141
Observations	109,276	109,276	35,626	35,626	58,695	58,695	13,605	13,605
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y
Firm FEs	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Table entries are the interaction terms of Equations (2.3) and (2.4). Dependent variable is log monthly wage. Sample is restricted to full-time employees up to 35 years of age. Observations denote individual by year observations. Sectors based on the 2003 edition of the German Classification of Economic Activities (WZ 2003). Manufacturing consists of codes D-G, service of K and L, public sector of N-Q, other sectors capture the residual sectors (agriculture, forestry, mining, hospitality industry and transportation). Standard errors (in parentheses) are clustered at the firm level. */**/** denote statistical significance at the 10/5/1% level, respectively.

For both firm types, the respective DID specifications—columns (3) and (5)—show a positive and statistically significant impact on female wages after the reform (the respective point estimates for new hires in the bottom panel are insignificant). Results from the TB model indicate that the overall pattern found in column (2) is determined to a larger extent by firms that pay above median wages: column (4) estimates that the annual convergence (divergence) between treatment and control group is zero, but that due to the reform there was a positive total average wage effect of 3.68 per cent for women over the 5-year post-reform period (5×1.02 - 1.42), statistically significant at the 5 per cent level. For firms paying below median wages—column (6)—the reform effect is less precisely estimated and indicates an overall effect of zero. Again, these empirical findings echo ex ante expectations: Section 2.2 has shown that the costs of maternity protection for the firm are higher for high-wage earners, and this feature of the regulation appears to have led firms to statistically discriminate against this group accordingly.

Additional results on wage effects by sectors are presented in Table 2.3, again distinguishing between all employees and new hires. Several patterns are worth noting:

first, estimation results from the DID model across sectors again point towards a relatively large—1.74 to 2.09 per cent—and statistically significant treatment effect on female wages, as indicated in columns (1), (3), and (5). Thus, the main finding of a positive reform effect wages of young women is not driven by a particular sector. Second, estimation results from the TB model indicate that the time trend of convergence (divergence) between male and female wages is estimated to be zero (columns (2), (4), (6), and (8)). The reform effect in the TB model by sector is relatively imprecisely estimated, and indicates a notable pattern only for the services sector, with an overall reform effect of 3.9 per cent over the 5-year post reform, significant at the 10 per cent level.

As sensitivity analysis we additionally included an interaction of the female binary variable and an indicator that equals one for the crisis years 2008-09. While the coefficient of this additional variable showed that men were hit harder than women by the crisis, the main result did not change. Moreover, the main result also proved robust to controlling for a level shift in female relative to male wages due to a reform of the parental leave system in 2007.

2.6 Conclusion

Despite substantial improvements in women’s labour market prospects in the past decades, women are still paid, on average, lower wages. Motherhood has been identified as one driving mechanism. While a large literature examines the *ex-post* career cost of motherhood, theory predicts that maternity may also be punished *ex-ante* through statistical discrimination by employers: specifically, negative effects on all women of childbearing age may arise if firms face direct costs of employing potential mothers. To test the hypothesis of statistical employer discrimination, we use the natural experiment of a maternity protection reform in Germany.

The empirical results contrast female and male wages within large firms and confirm the theoretical prediction: first, DID estimates show a statistically significant increase in female wages relative to male wages of 1.1 per cent in the post-reform period. This treatment effect estimate is in line with a back-of-the-envelope calculation, which puts the expected effect at around 0.8 per cent. Second, the trend-break model indicates that the general time trend in the convergence of female and male wages in large firms is not statistically different from zero, and the post-reform trend estimates a total reform effect of 2.5 per cent over five years. This result indicates that indeed under the pre-reform regulation employers showed statistically discriminating behaviour against potential mothers due to maternity protection wage continuation.

This evidence for statistical discrimination implies several policy conclusions. Since we find that labour market prospects improved significantly due to the reform, policy makers should try to identify factors that could result in statistical discrimination against potential mothers. A prominent example is the German parental leave legislation, where the current incentive structure leads women, who on average earn less than their partners, to take longer leave periods. Moreover, our findings support that it is worthwhile for the public to finance, through taxes, costs that occur asymmetrically to mothers only, if the aim is to prevent negative labour market effects *ex-ante* and *ex-post* for (potential) mothers, and to reduce gender-specific earnings differences.

CHAPTER 3

A Firm-Side Perspective on Parental Leave¹

3.1 Introduction

In imperfect labour markets, employment relationships generate rents for firms, making employment interruptions costly for them. One of the most important sources of employment interruptions are motherhood-related absences. While generous parental leave policies help parents to better reconcile work and family life after childbirth, firms need to handle the employment interruptions that these policies create. The costs of such interruptions may be substantial if the worker on leave cannot be easily replaced by other incumbent workers or external hirings. Such costs, in particular for small- and medium-sized firms, are a major reason brought forward in the US against a federal parental leave scheme (e.g., see Bartel et al., 2021). Yet, surprisingly little is known about the impact of parental leave absences on firms.

Our study helps fill this gap and improves our understanding of how firms deal with parental leave absences. We use the full population of administrative linked employer-employee data from Germany that allow us to uncover new stylised facts about firms' hiring behaviour as well as the relationship between parental leave absences and the availability of replacement workers inside and outside of the firm. We further examine how a paid parental leave reform granting additional parental leave benefits to medium- and high-earning mothers, incentivising them to delay their return to their employer in the first year after childbirth, affected mothers and their firms.

¹This chapter is joint work with Mathias Huebener (DIW Berlin and IZA), Daniel Kuehnle (Universität Duisburg-Essen and IZA) and Michael Oberfichtner (IAB). We are grateful for comments by Francine Blau, Anne A. Brenøe, David Card, Thomas Cornelissen, Max Deter, Bernd Fitzenberger, Christina Gathmann, Rita Ginja, Ines Helm, Simon Jäger, William Jergins, Brendon McConnell, Steven Rivkin, Kjell G. Salvanes, C. Katharina Spiess, Till von Wachter, and Matthias Westphal, as well as seminar and conference participants at ASSA 2021, DIW Berlin, COMPIE 2021, WEAI 2021, Hertie School of Governance, IAB, IAAE 2021, University of Erlangen–Nuremberg, Leuphana University of Lüneburg, Verein für Socialpolitik, and ifo Dresden. We thank Martin Popp for sharing data on employment by occupation at the district level.

Our paper makes three main contributions to the literature. First, we take a firm-side perspective on paid parental leave using rich employer-employee data that includes three crucial pieces of information: employment spells at the daily level, detailed occupational codes, and information on single locations of multi-site firms. This allows us to identify *local workgroups*, i.e., workers in the same occupation, same firm and same location, and separates us from related work as we are the only study to measure the number of substitute workers inside and outside of the firm for each mother. We provide first evidence concerning the exact timing and composition of replacement hirings, and we reveal the link between mothers' parental leave-taking and the availability of internal and external substitutes for two different parental leave systems. Focusing on a parental leave extension, we provide new evidence on the longer-term effects for mothers and their firms. As our analysis focuses on small firms with less than 50 employees, we contribute directly to the current debate about the implications of providing more generous parental leave schemes, as proposed by the Biden administration in April 2021 (NYT, 2021).

Second, we contribute new evidence on the substitutability of workers for temporary and agreed-upon absences. We focus on the case of parental leave absences, which rank among the most important reasons for worker absences as they directly affect the vast majority of women during their prime working lives.² The previous literature has so far focused on absences due to sickness and worker deaths. Parental leave absences, however, differ from absences due to sickness and worker deaths in important ways. Specifically, parental leave is typically anticipated and can thus be better planned by firms, and the length of parental leave is agreed upon in advance.

Our third contribution relates to statistical discrimination against young women in the labour market and the unintended consequences of public policies. We conjecture that profit-maximising firms anticipate and internalise the potential costs of longer birth-related absences by younger female workers and in turn reduce the hiring of younger women. To identify statistical discrimination net of other potentially confounding time trends, we apply an event-study framework inspired by Dobkin et al. (2018) and examine whether firms reduce their hiring of young women into occupations that *were not* affected by a birth in the same firm following the parental leave reform. If firms hire fewer women of childbearing age for workgroups with few internal substitutes after the reform, such patterns would be evidence for statistical discrimination. Despite its theoretical appeal, identifying statistical discrimination is notoriously difficult and the existing evidence mostly stems from audit studies.

²In most OECD countries, 80-90 percent of women have at least one child before age 45 (see UN World Fertility Data).

In a first step of the analysis, we examine how firms adjust their hirings and separations to birth-related worker absences and how workplace characteristics are linked to mothers' length of parental leave. We find that replacement hiring is more pronounced when few internal substitutes, i.e., workers in the same occupation, are available, whereas external substitutes, i.e., the share of workers in the same occupation in the local labour market, are less important. We find no evidence for an adjustment in separations to reduce the impact of the absent mother. We then show that mothers substantially postpone their return-to-work if more internal substitutes are available in their firm, whereas—as for replacement hiring—external substitutes are not related to the length of parental leave. These links hold when we further control for maternal characteristics, and occupational and regional heterogeneity. We then investigate how these patterns change with the introduction of a more generous paid parental leave scheme which incentivises longer absences from work during the first year after childbirth. The reform almost eradicates the link between the length of parental leave and the availability of internal substitutes during the period of benefit receipt. These results suggest that the introduction of paid parental leave may distort the coordination between employers and mothers in the benefit payment period.

In a second step, we study the effects of the parental leave expansion on mothers' return to their pre-birth employer and on firms. Our main firm outcome is the employment level: As employment generates rents for firms, lower employment implies lower profits, *ceteris paribus*. Our empirical estimation strategy employs a dynamic difference-in-differences design. In line with the economic incentives, we find that medium- and high-earning mothers substantially delay their return to their pre-birth employers when parental leave benefits are expanded, with no medium- to longer-term impact on the probability of exiting their firm. Mothers giving birth after the reform have a 20 percentage points lower probability to have returned to their pre-birth firm ten months after childbirth. We find negative effects on firms' employment which implies that firms do not fully compensate the longer absences of mothers. The effects are driven by firms in which fewer internal substitutes are available for the mother-on-leave. As the number of internal substitutes increases, the employment gap decreases and eventually disappears. We observe that firms with more internal substitutes experience a generally higher labour turnover that may help to bridge the labour shortage. We do not find evidence that firms hire more workers from external labour markets to compensate the longer worker absence. Despite the short-term labour gap in firms with few internal substitutes, we identify no effects on firms' employment in the second to fourth year after childbirth.

As firms with fewer internal substitutes cannot fully compensate the short-run negative labour supply shock of longer absences, we conjecture that they may internalise the associated costs in their future hiring decisions and analyse whether firms' hiring under the new policy regime deviates from their pre-reform hiring history in terms of observable worker characteristics. We find that firms are less likely to hire women of childbearing age after the reform into occupations in which only few internal substitutes were available in case of a pregnancy. In contrast, the hiring of older women and men into those occupations increases.

Our paper ties into three strands of economic literature. First, our paper adds to a small literature on the effects of parental leave for firms. A recent paper by Ginja et al. (2020b) studies a parental leave expansion from 12 to 15 months in Sweden and finds that the reform increased separations of mothers from their pre-birth firms and that adjustments to the longer absences are costly for firms. We complement their study by focusing on a reform that (i) substantially expanded parental leave absences *within* the first year after childbirth and (ii) allowed firms to partly anticipate the longer leave absences which could impact firms' replacement hiring decisions. Related, Gallen (2019) studies the effects of a Danish parental leave reform on firms and mothers' coworkers. The reform increased the length of fully-compensated parental leave by 22 weeks within the first year after childbirth. She finds decreases in mothers' retention probability and in firms' survival rates. For Denmark, Brenøe et al. (2020) use Danish administrative data and examine the joint effect of motherhood and the subsequent parental leave period on the mother's firm by combining a matching- and event-study-approach; their findings suggest that the costs of parental leave on firms are negligible. For the US, Bartel et al. (2021) survey small- and medium-sized firms to study the introduction of a paid parental leave policy in New York covering eight weeks of partial wage replacements. They find a short-term increase in employers' self-reported ease of handling employee absences for firms with 50 or more employees, but not for smaller firms.³

Second, we contribute to the literature on the substitutability of workers, with a particular focus on firm-specific capital and firms' ability to substitute (temporary) absences of workers. Previous studies have focused on sickness absences (e.g., Hensvik and Rosenqvist, 2019) and worker deaths (e.g., Jäger and Heining, 2019).⁴ Specifically,

³Our paper also relate to the much larger literature on the effects of parental leave policies on maternal labour market outcomes (e.g., see Lalive and Zweimüller, 2009; Schönberg and Ludsteck, 2014; Rossin-Slater, 2018; Kleven et al., 2020).

⁴One strand of the literature on the substitutability of workers focuses on how the death of key figures within firms, such as CEOs, superstar scientists, or inventors, affect the productivity and earnings of their coworkers (Azoulay et al., 2010; Jaravel et al., 2018; Bennedsen et al., 2020). These results cannot immediately be generalised about the substitutability of "regular" employees since this

Hensvik and Rosenqvist (2019) use Swedish administrative data and show that firms keep sickness absences low for positions where workers are harder to replace. Jäger and Heining (2019) document that the unexpected death of a worker affects their coworkers' earnings and firm retention probability using German administrative data. They find that the effect depends on the degree of substitutability between workers and on the availability of external substitutes in the labour market.

Third, we add to the quasi-experimental literature on statistical discrimination against young women in the labour market (Fernández-Kranz and Rodríguez-Planas, 2021; Jessen et al., 2019). Child-related work absences require firms to find a replacement during leave, affect the accumulation of firm-specific human capital, and may deteriorate the skills of employees during leave (Adda et al., 2017). If those absences are costly for firms, theory predicts that employers would internalise the associated costs and discriminate against women of childbearing age by hiring or promoting them with a lower probability or by paying them lower wages than their male counterparts. An extension of paid parental leave that is meant to facilitate the reconciliation of work and family life, incentivises women to return to work later. Our study provides new evidence that such family policies can potentially backfire on potential mothers in the labour market.

The remainder of this paper is structured as follows. We provide information on the background and institutional details in section 3.2 and describe the employer-employee matched data and our sample in detail in section 3.3. Section 3.4 describes how firms' hirings and separations respond to a birth and establishes basic relationships between parental leave schemes, mothers' absences and workplace and labour market characteristics in different parental leave regimes. In section 3.5, we present the parental leave reform effects on workers and firms. In section 3.6, we evaluate the impact of the parental leave expansion on firms' longer-term hiring to assess whether parental leave expansions can result in statistical discrimination against women of childbearing age. Section 3.7 concludes and discusses some policy implications.

3.2 Background and Institutional Environment

3.2.1 Relevant Aspects of the German labour market

Our analysis focuses on Germany which for a long time has been characterized by low labour force participation rates of mothers with dependent children. However, a series of policy reforms starting in 1996 substantially expanded the availability of day care

literature examines the loss of very specific workers characterised by exceptional productivity and/or firm-specific capital.

and strongly encouraged early maternal employment. In 2006, the day care attendance rate for 0-2 year olds, including centre-based and family day care services, was 13.6 percent (cf. OECD average 30 percent, based on OECD, 2016a,b) and 41 percent of these mothers were employed (BMFSFJ, 2018). 63 percent of women aged 25-54 with at least one child aged 0-14 participated in the labour force (cf. OECD average 66.1 percent).

One key feature of the German labour market relevant for our analysis is the important role played by small firms, on which we focus in our analysis. Small firms are defined as firms that employ less than 50 workers and earn less than 10 million euro in annual turnover. These firms represent a large share of the German labour market as they make up 96.9 percent of all enterprises and employ about 41 percent of all workers (Destatis, 2018). As the debate about possible adverse effects of parental leave on firms is centered around small and medium-sized firms, Germany provides a suitable setting to study this question.

Although the extent of temporary work agencies has increased in many countries over the past twenty years, temporary employment was of little importance for the firms we analyse (less than 50 employees) during the period we study. Specifically, between 2003 and 2010, the percentage of temporary workers among all employees in Germany was less than 0.7 percent in firms with 10-49 employees and less than 0.4 percent in firms with less than 10 employees (for more details, see Hirsch and Mueller, 2012). Workers hired from temporary work agencies therefore only play a minor role to substitute workers on parental leave in our setting.

3.2.2 Family Policies Supporting Women in the Labour Market

German family policy supports pregnant women and mothers in the labour market through the following key policy measures, which are also relevant for their employers: paid maternity leave, job protection, and parental leave benefits.⁵

Job Protection and Parental Leave. Parents can claim job-protected parental leave (*Elternzeit*) from their employer allowing them to return to their previous position within 36 months after childbirth. To claim job-protected parental leave subsequent to maternity leave, mothers have to notify their employer at the latest one week after childbirth. The period for which parental leave is claimed is then binding. While on job-protected leave, parents are allowed to work part-time.

Paid Maternity Leave. All mothers are entitled to paid maternity leave which lasts from six weeks before expected delivery to eight weeks after childbirth. During this

⁵Additionally, day care spots are publicly subsidised and not-for-profit, as only three percent of institutions are run by private and non-charitable organisations.

period, (expecting) mothers are generally not allowed to work but they are entitled to a full (net) earnings replacement.⁶ The statutory health insurance companies pay for the earnings replacements, so that firms do not incur any direct costs.⁷

Parental Leave Benefits. Parental leave benefits are an important determinant of the length of parental leave (e.g., see Schönberg and Ludsteck, 2014). In Germany, parental leave benefits were substantially reformed in 2007 from a means-tested to an earnings-based scheme. In the following, we describe both policy regimes as we exploit the reform in our analysis.

Prior to the reform in 2007, parents with low household income were eligible to receive benefits for up to 24 months after childbirth. These publicly-funded benefits were means-tested and paid a maximum of 300 euro per month (about 370 USD in 2006), corresponding to around eleven percent of average pre-birth net household income. Families qualified for the benefit if their annual net income was below a certain threshold, which varied with the household structure, number of children, and time since giving birth.⁸ About 77 percent of parents were eligible to receive benefits for up to six months after childbirth (see Huebener et al., 2019). Due to repeated means-testing and lower household income thresholds for eligibility, the share of eligible parents fell to 47 percent for 7-12 months after childbirth and to 40 percent for 12 to 24 months after childbirth.⁹

The 2007-Paid Parental Leave Reform. In September 2006, the German government substantially reformed the paid parental leave system which affected parents of all children born on or after January 1, 2007. The new benefit system replaced the previous means-tested benefits with an earnings-based paid parental leave system. Benefits were

⁶Mothers are allowed to work up to childbirth if they provide their explicit consent and if the work environment is considered safe. After childbirth, they are not allowed to work for the duration of maternity leave (see §3 *Mutterschutzgesetz*).

⁷Administratively, both health insurers and firms make earnings continuation payments. However, firms' expenses are fully reimbursed by health insurers. Prior to 2006, for firms with more than 30 employees, health insurers covered only 13 euro per day, and firms had to cover the excess amount of earnings continuation payments to cover women's pre-birth net earnings (Jessen et al., 2019). Changes in these regulations are not a threat to our quasi-experimental analysis where we focus on a paid parental leave reform in 2007 and analyse births after this policy change.

⁸For the first child, it was possible to receive these benefits for six months after childbirth if the net yearly income of a couple was less than 30,000 euro (23,000 euro for single parents). From the seventh month onward, this limit was 16,500 euro for couples (13,500 euro for single parents). If the expected net income exceeded these thresholds, the parental benefits were reduced rather than entirely withdrawn. The reduced child-rearing allowance was only available for families with an expected net income of up to 22,086 euro for couples (up to 19,086 euro for single parents; for details, see BMFSFJ, 2008).

⁹Part-time work of up to 30 hours per week was permitted during the benefit receipt period. Parents eligible for benefits for up to 24 months could also choose higher benefits (450 euro) for up to 12 months. For children born in 2005 and 2006, only ten percent of all parents chose this option (own calculations based on SOEPv30).

paid for up to 12 months to either parent and replaced 67 percent of the average net labour income earned in the 12 months prior to childbirth.¹⁰ The benefit had a floor of 300 euro and was capped at 1,800 euro per month.¹¹ Given near universal eligibility, take-up was almost 100 percent (Destatis, 2008).

The reform did not change the maternity leave period, the 36-months job protection period, or part-time employment regulations during the job-protected period. Overall, the direct costs of childbirth-related absences for employers are negligible, the mostly indirect costs relate to finding suitable replacement workers and to consequences on firm's operating processes.

Our analysis focuses on medium- and high-earning mothers who unambiguously gained higher paid parental leave benefits following the two months of mandatory maternity leave. For example, a mother earning 2,500 euro (net) per month before childbirth could claim 1,675 euro per month, or 16,750 euro in total, after the reform. In comparison to the maximum regular benefit of 300 euro under the old benefit system, the reform clearly reduced the opportunity costs of longer workplace absences for medium- and high-earning mothers.

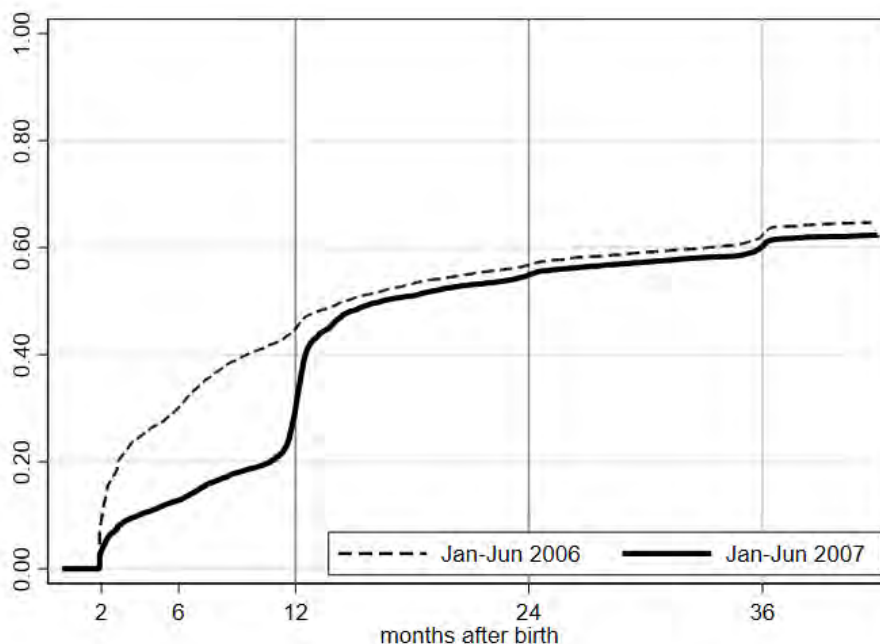
To illustrate the reform effect on maternal employment, Figure 3.1 plots the share of mothers who have returned to their pre-birth employer at different points in time, distinguishing between mothers who give birth in the same calendar months (January to June) before and after the paid parental leave expansion. With lower benefits (dashed line), mothers return gradually after the end of their maternity leave period. With extended paid parental leave (solid line), the return within the first year is substantially delayed and starts to converge to the pre-reform pattern with the expiry of the parental leave benefits after 12 months. This pattern is consistent with the changed economic incentives during the first year.¹²

¹⁰Two additional months were granted for single parents or if both partners take parental leave for at least two months. The maximum length of 14 months of paid parental leave could be split flexibly between both parents, with a minimum of two months per parent. Approximately 96 percent of parents assign the main benefit period (>7 months) to the mother. In our observation period, 15 percent of fathers take paid parental leave, mostly for two months. Parents could also choose to receive only half of the monthly benefits to double the benefit period, with only eight percent of parents choosing this option (Destatis, 2008).

¹¹Individuals who did not work prior to giving birth, or those with low earnings, continue to receive 300 euro per month, but now only for up to 12 months instead of 24 months.

¹²In Appendix Figure 3.A.1 we show that the pattern is similar when we examine the share of mothers who have returned to any employment.

Figure 3.1: Return to pre-birth firm by birth semester



Notes: Figure shows the share of mothers with pre-birth earnings of at least 1700 euro per month who have returned to their pre-birth employer by month t after childbirth. The dashed line indicates mothers giving birth between January and June 2006 (pre-reform), the solid line indicates mothers giving birth between January and June 2007 (post-reform). Source: IEB, own calculations.

3.3 Data

3.3.1 Data Source: Social Security Records

We use administrative data from Germany that cover the universe of firms and workers subject to social security contributions (the IAB Integrated Employment Biographies, IEB V13.00.00, see Jacobebbinghaus and Seth, 2007, for a detailed description). The data is available from 1975 through 2018 and cover about 82 percent of all workers in Germany.¹³ The reported earnings and job durations are used to calculate social security payments and benefits and are therefore highly reliable.

Several features of the data render it particularly suitable for our analysis. The first advantage is that the data contain information on employment spells at the daily level as employers report the precise start and end dates of any employment spell.

¹³Civil servants and self-employed individuals are not included in the data. This implies that we do not have information on all workers in the public sector who are subject to social security; we therefore exclude the public sector from our analysis. The lack of self-employment spells is not a problem for our analysis, as the main units of analysis are the firm and the workplace. Any parental leave effects on selection into self-employment or the public sector would only affect the return to the same firm that we can fully observe.

This allows us to exploit the exact timing when new employment spells start and end. Second, the data allow us to identify single locations of multi-site firms. For simplicity, we refer to these establishments as firms throughout the paper. Third, we have detailed occupational information for workers at the 3-digit level (with 256 unique occupations of mothers in our sample). Combined with the exact location, the data allow us to identify *local workgroups*, i.e., workers in the same occupation, same firm and same location.

In addition to the above features, the data furthermore include basic socio-demographic characteristics like workers' gender, citizenship, education (imputed as described in Thomsen et al., 2018) and date of birth. As the data goes back to 1975, we can reconstruct the entire employment biographies subject to social security contributions of all individuals in our sample. The data also include a part-time/full-time indicator, but no further details on working hours. The data do not include direct information on motherhood, but we follow Müller and Strauch (2017) to identify mothers and infer the expected date of delivery by exploiting the legal requirement that employers have to notify the health insurance companies, who carry out the maternity leave payments, about the start date of this leave period.¹⁴ We use the expected date of delivery to assign mothers to specific paid parental leave regimes. To avoid the misassignment of births around the policy cut-off, we exclude all expected births that occur two weeks before and after January 1st from the analysis.

3.3.2 Operationalising Internal and External Substitutes

We define workers as *internal substitutes* if they work in the same firm and in the same 3-digit occupation (i.e., they perform similar tasks) ten months prior to childbirth. This coworker definition is also used in Cornelissen et al. (2017) and Hensvik and Rosenqvist (2019). Throughout the paper, we refer to mothers' coworkers as *internal substitutes* and we use the term *workgroup* when we additionally include the mother. We define three groups which correspond to terciles: mothers with 0-1, 2-5, and 6 or more internal substitutes.¹⁵

To measure the availability of *external substitutes*, we build on the concept of labour market thickness. From a worker's perspective, a market is thick when she receives many offers for a given amount of search effort (Lazear, 2009). From a firm's per-

¹⁴The same notification code is used in some cases of longer illnesses, but this event is very rare for women of childbearing age and such absences are often shorter than the required mandatory leave lengths for childbirth. We implement several checks to ensure that the notification reflects childbirths.

¹⁵Appendix Figure 3.A.2 shows the distribution of 1-digit occupations following the classification by Blossfeld (1987).

spective, a market is thick if the frequency of receiving suitable applicants for a given vacancy is high. As an empirical proxy for external substitutability, we follow Jäger and Heining (2019) and use the relative density of 3-digit occupational groups in the local labour market of a firm.¹⁶ For this purpose, we calculate the regional share of employment in each occupation relative to the nationwide share of employment in each occupation as of June 30, 2006 from the universe of the IEB. We also split labour market thickness as a measure for external substitutes into terciles.

3.3.3 Sample Selection and Treatment Assignment

In our setting, a workplace is affected by the 2007 parental leave reform if a mother employed by the firm gives birth on or after January 1, 2007. As the date of birth cut-off determines the paid parental leave eligibility, this institutional rule assigns mothers and their firms into a natural treatment group (births between January and June 2007) and control group (births between July and December 2006). To account for any seasonality in outcomes, we further consider mothers and firms with births in the preceding year (July 2005 to June 2006) within our estimation strategy.

We focus our analysis on *private, for-profit* establishments and drop establishments that are part of the government, military, churches and other non-profits, as their substitution and wage setting processes substantially differ from private sector firms. As we expect effects to be concentrated in smaller firms, we focus on firms with up to 50 employees before the pregnancy occurs in the firm. The mean firm size in our analysis sample is 14 employees with a median size of ten (see Appendix Figure 3.A.3).

As firms could experience multiple births – before and after the reform – we focus on firms in which exactly one birth occurs in the period between two years before and two years after childbirth. This sampling restriction allows us to cleanly identify firms as being affected by the paid parental leave reform and to contrast them with firms that were not affected. As we already focus on smaller firms, considering only one birth per firm is not a restrictive condition.

Further, we only keep firms where the mother giving birth fulfils the following three criteria. *First*, we focus on first-time mothers as these are more strongly attached to the labour market. We would therefore expect that the effects of a birth and of parental leave are more pronounced compared to mothers with higher-order births.¹⁷ *Second*, we

¹⁶Our classification of labour market regions follows Kosfeld and Werner (2012) who define 141 regions in Germany based on commuting flows.

¹⁷We also do not focus on higher-order births as these can only be identified in the data if the mother returns to work between two births. Including mothers with higher-order births could thus yield a selective sample with respect to birth-spacing and mothers' labour force attachment, especially if the parental leave reform affects these outcomes.

only keep mothers with gross monthly earnings of at least 1700 euro before giving birth (this excludes the lower 37 earning percentiles of first-time mothers). The introduction of paid parental leave unambiguously increased non-labour income for these mothers during the first year after giving birth, thus monotonically increasing mothers' financial incentives for longer absences from work.¹⁸ *Third*, we impose a minimum tenure requirement and focus on mothers who have been at their workplace for at least ten months prior to giving birth. This restriction avoids endogenous selection into firms and occupations during pregnancy.

Our final analysis sample contains 23,679 mothers and firms. We observe the entire employment history of all workers who have been employed at those firms at any time from four years before to eight years after birth. Table 3.1 reports pre-birth characteristics for mothers and their firms. Column (1) shows that mothers in our data are on average 30 years old at childbirth, 96 percent are German citizens. Around 39 percent of mothers have high levels of education and their average gross annual earnings amount to around 31,000 euro (38,000 USD in 2006). These mothers are strongly attached to the labour market as around 94 percent worked full-time before childbirth with an average firm tenure of just under five years. 90 percent of firms are located in West Germany. Average firm size amounts to 14 employees, and mothers are in workgroups with close to six workers (i.e., five internal substitutes), and just below two-thirds of employees in the firm are women. Most firms are in the service sector.

To assess how the selection criteria affect our sample, Appendix Table 3.A.1 compares the characteristics of mothers in our analysis sample with all excluded first-time mothers who gave birth in the two-year sample period. In particular, mothers in our sample are older at birth (30 vs. 28.6 years), have obtained higher education (39 percent vs. 31 percent), have higher monthly pre-birth earnings (2664 euro vs. 2125 euro), higher firm tenure (4.65 years vs. 3.8 years), and are more likely to work full-time pre-birth (94 percent vs. 82 percent). Consistent with the above differences, mothers in our analysis sample are more strongly attached to the labour market as reflected by the slightly higher shares of mothers who return to the labour market within one and three years after childbirth.

With respect to mothers' firms, Appendix Table 3.A.2 shows that firms included in our sample are substantially smaller compared to the excluded firms (14.5 vs. 90

¹⁸In contrast, the 2007 paid parental leave reform reduced some low-income mothers' family non-labour income in the second year after childbirth (see section 3.2.1). The reform might hence induce some low-earning mothers to return to work earlier and others to return later (Kluve and Schmitz, 2018; Huebener et al., 2019). We thus cannot exploit the 2007 paid parental leave reform to investigate the effects of longer leave on those mothers and their employers.

Table 3.1: Summary statistics and balancing

	Sample window					DD coef.
	All	Jul-Dec 05	Jan-Jun 06	Jul-Dec 06	Jan-Jun 07	
	(1)	(2)	(3)	(4)	(5)	
<i>Individual mother characteristics</i>						
Age in years	29.963	29.788	30.135	29.845	30.112	-0.079 (0.104)
German citizenship	0.958	0.960	0.954	0.962	0.957	0.001 (0.005)
High education	0.390	0.373	0.390	0.389	0.410	0.003 (0.013)
Annual earnings in year before birth	30539.930	31286.521	29951.461	31118.018	29674.762	-108.195 (269.211)
Tenure at current firm in years	4.648	4.613	4.581	4.716	4.683	-0.002 (0.098)
Full-time employed	0.940	0.943	0.942	0.938	0.937	0.000 (0.006)
<i>Pre-birth firm characteristics</i>						
Location in West Germany	0.895	0.900	0.893	0.894	0.893	0.006 (0.008)
Firm size	14.062	14.218	13.983	14.047	13.984	0.172 (0.298)
Workgroup size	5.733	5.828	5.756	5.685	5.652	0.040 (0.165)
Share of female employees	0.631	0.629	0.629	0.632	0.634	0.002 (0.007)
<i>Sector</i>						
Agriculture, fishing and mining	0.013	0.013	0.013	0.012	0.013	0.001 (0.003)
Manufacturing	0.125	0.127	0.121	0.125	0.126	0.006 (0.009)
Electricity, gas, water	0.004	0.003	0.005	0.002	0.003	-0.001 (0.002)
Construction	0.041	0.039	0.042	0.045	0.036	-0.013** (0.005)
Wholesale and retail	0.332	0.321	0.333	0.339	0.334	-0.017 (0.012)
Hotels and restaurants	0.020	0.021	0.019	0.019	0.020	0.003 (0.004)
Transport, storage, communication	0.051	0.047	0.054	0.054	0.051	-0.011* (0.006)
Financial intermediation	0.065	0.059	0.070	0.065	0.068	-0.008 (0.006)
Real estate, renting, business activities	0.304	0.297	0.295	0.306	0.321	0.017 (0.012)
Observations	23,679	6,360	5,680	6,003	5,636	23,679

Notes: Table shows pre-determined characteristics at the individual level of the mother and at her pre-birth firm. Mean values are presented in columns (1)-(5). The coefficients in column (6) are obtained from a difference-in-differences specification. Robust standard errors in parentheses. * < 10% ** < 5% *** < 1%

Source: IEB, own calculations.

employees in June 2006).¹⁹ Firms in our sample are more likely to come from West Germany (90 percent vs. 82 percent), have slightly older employees (38.6 vs. 37.3 years), and pay higher median gross wages (2564 euro vs. 2174 euro). With respect to the industry structure, we observe some small shifts, in particular that firms in our

¹⁹In the main table, this number is calculated ten months pre-birth, in Appendix Table 3.A.2 we get slightly different numbers as we use information from the Establishment History Panel (BHP) which is based on June 30, 2006, as our data does not cover all employees from dropped firms.

sample are less likely to come from manufacturing, and hospitality, and more likely to come from other services.

One potential concern that emerges from selecting firms with one birth during a four-year period is that the reform effects on fertility may cause an endogenous sample selection bias. For example, if women were more likely to give births after the reform (in the medium-run), we would be more likely to exclude firms with more women of childbearing age. If this was an issue for our analysis, we would expect systematic differences in pre-determined observable characteristics between treatment firms and control firms, as treated firms were exposed longer to the new parental leave regulation during the period considered for our sample selection. However, we find no evidence for such systematic differences between treatment and control firms (see column (6) of Table 3.1 for balancing checks, and section 3.5.1 for further details).

3.3.4 Outcome Variables

We first study how the reform affects mothers' return to their pre-birth firm to characterise the employment gaps that an increase in parental leave causes. We take advantage of the daily level of the employment data and define binary indicators for mothers working at their pre-birth firm at the monthly level, allowing us to trace out the prolonged absence of mothers in detail. As earnings in the data are reported as a daily average over the reporting period of the employment spell (at most one calendar year), we cannot reliably calculate monthly earnings. Instead, we compute the annual earnings of mothers at their pre-birth firms and deflate earnings to a common base CPI of 2010.²⁰

To measure firm outcomes, we use firms' employment level and their total wage bill, based on the following considerations: In labour markets with imperfect competition, employment generates a surplus that accrues partly to the worker and partly to the firm (for an overview, see Manning, 2011). *Ceteris paribus*, most importantly holding constant other inputs and the production technology, lower employment hence implies lower profits. We therefore use employment as our main firm-level outcome and measure it as the number of workers at a firm. We analyse employment—as for mothers—at the monthly level.²¹ Because we do not have precise information on hours worked, we additionally examine firms' wage bill. If the number of hours worked by each worker increases in response to mothers' absences (an intensive margin response), the wage bill

²⁰Earnings are top-coded at the social security contribution ceiling, which affects less than one percent of mothers in our analysis sample and less than 2.5 percent of their coworkers. Top-coded earnings are assigned the coding-threshold value, i.e., we cannot capture effects above the earnings maximum.

²¹The data allow for an analysis at the daily level, but we use a monthly aggregation to reduce computational demand.

should respond less than total employment to the absence. The wage bill is however also affected by any wage changes necessary to increase other workers' labour supply, e.g., overtime pay, and the interpretation is therefore less clear than for employment. Analogous to mothers' earnings, we measure the wage bill of the firm at the annual level.²² The aggregated wage bill serves as a proxy for output of the firm. To make the estimations comparable across firms of different size and across the outcomes, we consider all firm-level outcomes relative to the baseline period. We furthermore censor firm outcomes at the 99th percentile of the respective distribution to reduce the imprecision induced by outliers.

3.4 Firms and Parental Leave Absences

3.4.1 Replacement Hirings and Separations

Firms can either hire replacements or reduce separations to substitute absent mothers. We start with replacement hirings and consider hirings in the same occupation as the mother as potential replacements. To examine firm's behaviour in a stable institutional setting, we focus on births occurring prior to the reform in 2006.²³

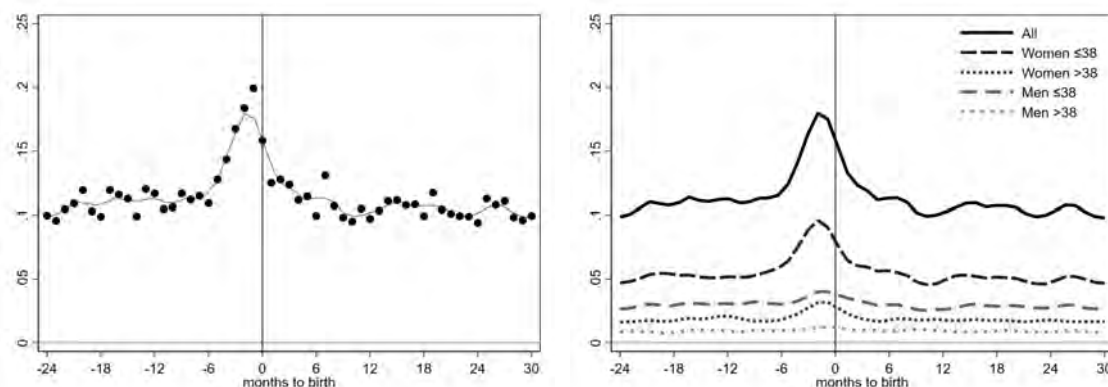
Panel A of Figure 3.2 plots the monthly number of hirings in mothers' workgroups from 24 months before birth up to 30 months after birth. Until six months prior to childbirth, firms constantly hire around 0.1 workers per month. Six months before childbirth, hirings start to gradually increase. This coincides with the end of the first trimester, at which a pregnancy is typically considered safe and when mothers usually announce their pregnancy to their employers. The observed pattern shows that firms hire replacement workers from external labour markets and also allow for some transitional period before mothers go on leave, most likely to hand over tasks to ensure a smooth transition. We observe 0.28 excess hirings in the six months to birth compared to the same months one year before, i.e., not all mothers are replaced by firms (similarly, Jäger and Heining, 2019, find that less than half of deceased workers are replaced). In the period following women's childbirth, the average hiring rate returns to the pre-birth level.²⁴

²²In contrast to the the Danish setting analysed in Brenøe et al. (2020), firms in Germany are not responsible for carrying out paid parental leave payments to mothers and this is accordingly not reflected in firms' wage bill.

²³In Appendix Figures 3.A.4-3.A.6 we show the same figures for the post-reform period in 2007 with very similar patterns.

²⁴In Appendix Figure 3.A.7 we present hirings in all *other* occupations in the firm. While a small increase in hiring around childbirth is observed, pointing to cascade effects in the firm or some replacement hiring in other occupations, the increase is smaller in both absolute and relative terms, thus supporting our definition of internal substitutes. This is in line with Jäger and Heining (2019)

Figure 3.2: Hirings around childbirth



A: Replacement hiring around childbirth

B: Composition of replacement hirings

Notes: Figure shows hiring in the same occupation of mothers around childbirth. Mothers returning to their pre-birth workgroup are not counted as hirings. Numbers of hirings are cleansed of calendar month effects. Sample period is births from January 2006 to June 2006. Source: IEB, own calculations.

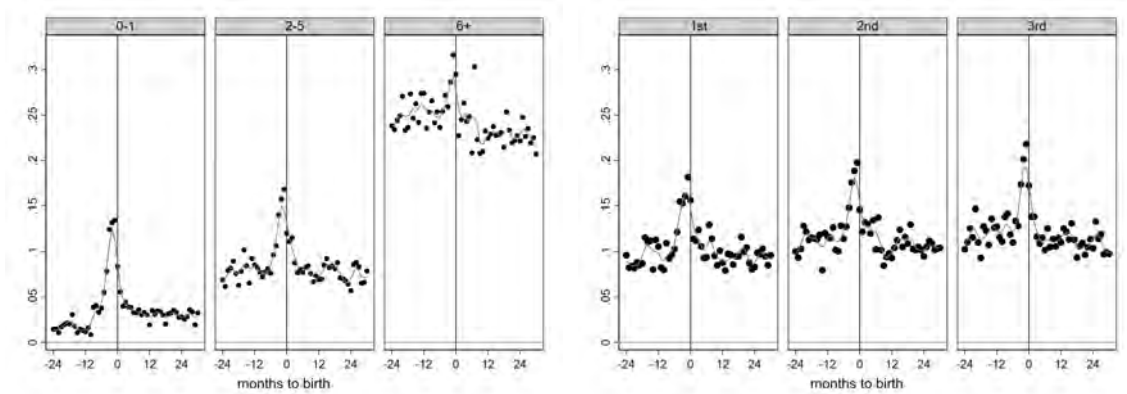
To examine the composition of the additional hirings, Panel B plots the evolution of hirings for four mutually exclusive demographic groups: men and women, above and below age 38. Almost half of replacement hiring comes from the same demographic group as mothers, i.e., young women, which corresponds to their share among hirings earlier and later. The relative uptick in hirings around childbirth is also stronger for older women than for men, suggesting that future mothers are more frequently replaced by women of all ages rather than by men.

Figure 3.3 investigates how firms' replacement hirings differ by the availability of internal and external substitutes. According to Panel A, the peak in the months leading up to childbirth is most pronounced when fewer internal substitutes are available. With more internal substitutes, the level of hirings is generally higher, in line with normal churning, and the replacement hirings are less pronounced. Excess hirings in small workgroups amount to 0.34 (0-1 substitutes) and 0.33 (2-5), but with more substitutes (6+) only 0.14 excess hirings occur. Panel B provides no evidence for substantial differences in replacement hirings between thick and thin labour markets. Thus, the availability of external substitutes does not appear to be a main hindrance for firms when trying to replace mothers.

Alternatively, firms may reduce separations of incumbent workers to keep the employment stock constant. We examine this in Panel A of Figure 3.4, where we plot the separations of incumbent workers from ten months prior to childbirth up to 30 months

who show that around three-quarters of hirings in respond to the death of a worker occur in the same occupation.

Figure 3.3: Hirings by availability of substitutes

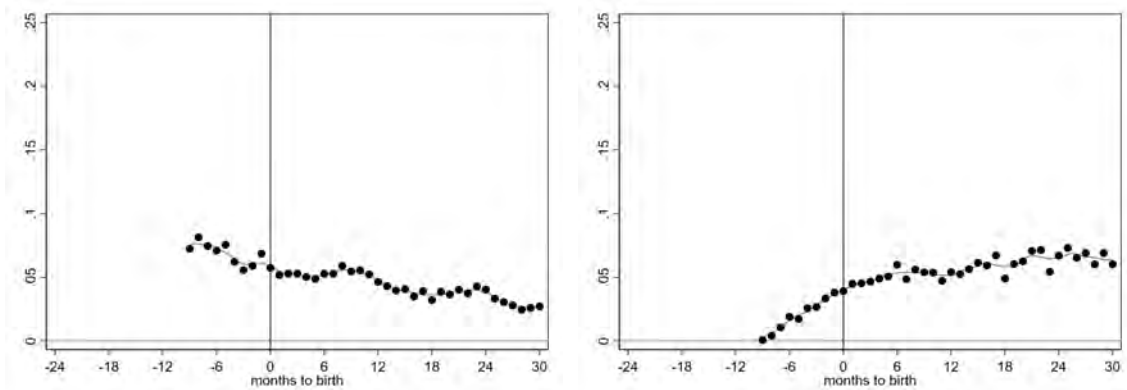


A: Number of internal substitutes (terciles) B: External substitutes (thickness terciles)

Notes: Figure shows residualised hiring by availability of internal and external substitutes. Internal substitutes are defined as the number of coworkers in the same occupation ten months pre-birth. External substitutes distinguished by terciles of labour market thickness of the occupation. Sample period is births from January 2006 to June 2006.
Source: IEB, own calculations.

after childbirth. The figure shows that separations decline smoothly over time. In Panel B of Figure 3.4, we examine separations of non-incumbent workers, i.e., workers newly hired into the firm starting ten months prior to childbirth. Here, we observe a steady increase in separations. Overall, the figure shows that separations are not reduced to substitute absent workers.

Figure 3.4: Separations by incumbents and replacements



A: Incumbents

B: Replacements

Notes: Figure shows residualised separations of workers in the same occupation as mothers. Panel A shows separation events for incumbents, defines as coworkers ten months before birth. Replacements in Panel B are those hired afterwards. Sample period is births from January 2006 to June 2006.
Source: IEB, own calculations.

Given the relationship between the availability of substitutes and firms' replacement hirings, we next examine how the length of parental leave depends on the availability of internal and external substitutes in different institutional settings.

3.4.2 Mothers' Return Behaviour with Short Paid Parental Leave Benefits

We start in a policy environment with limited paid parental leave benefits, where all mothers receive full earnings replacements for the first two months after childbirth, but only some mothers receive a small amount of parental leave benefits three to six months after childbirth (i.e., births between January and June 2006, see section 3.2 for details).²⁵ We would expect that mothers with fewer internal substitutes return earlier to their pre-birth employer than mothers with more substitutes.

Panel A of Figure 3.5 provides the Kaplan–Meier plot for returning to the pre-birth employer, separately by the number of internal substitutes. The key finding that emerges is that parental leave is substantially shorter if only few internal substitutes are available. In particular, mothers with 0-1 substitutes return the earliest, while mothers with six or more substitutes are the slowest to return. Mothers in between these groups initially return like mothers with many substitutes, but then move towards mothers with few substitutes. All groups converge towards the child's third birthday, which coincides with the end of the job protection period. Panel B of Figure 3.5 shows return by the availability of external substitutes in the respective labour market region, split into terciles. In the first year after childbirth, mothers' return to the pre-birth employer is very similar across thin and thick labour markets. We then observe that mothers in thinner labour markets return slightly earlier than mothers in thicker labour markets in the second and third year after childbirth.

Given that workers' characteristics could correlate with the number of internal and external substitutes, the observed patterns could be driven by differences in worker, occupation or firm characteristics.²⁶

To account for these differences, we estimate the following regression model:

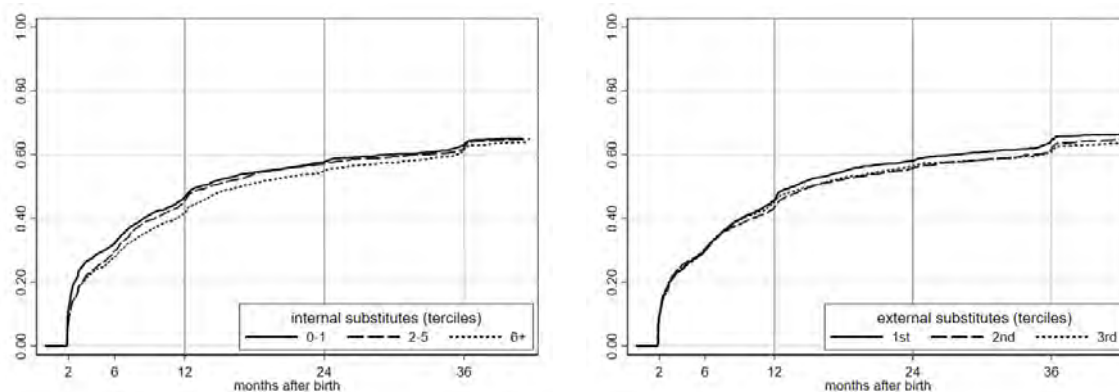
$$\begin{aligned} return_i^t = & \alpha^t + \delta^t \ln(\text{internal substitutes})_i \\ & + \gamma^t \ln(\text{external substitutes})_i + X_i' \beta^t + \epsilon_i^t \end{aligned} \quad (3.1)$$

where $return_i^t$ is a binary indicator for mother i to have returned to her pre-birth employer t months after giving birth where $t \in (3, \dots, 42)$. We use this binary outcome

²⁵We use births from January to June 2006 to be able to directly compare these later to births from January to June 2007 without potentially distorting calendar month effects.

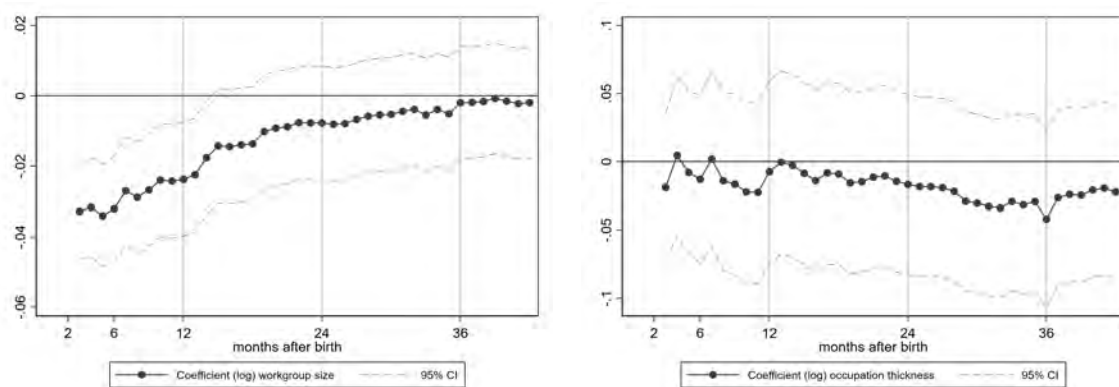
²⁶See Appendix Table 3.A.3 for summary statistics which differentiate between different levels of internal substitutability and Appendix Table 3.A.4 which shows that individual and firm characteristics are balanced within those subsamples as well.

Figure 3.5: Return to pre-birth firm in setting with low paid leave entitlements



A: Kaplan-Meier by internal substitutes

B: Kaplan-Meier by external substitutes



C: Internal substitutes (estimates)

D: External substitutes (estimates)

Notes: Sample period is births from January 2006 to June 2006. Panels A and B show Kaplan-Meier failure functions of mother's return to their pre-birth firm. Panel A differentiates by the number of internal substitutes at her firm defined as coworkers in the same occupation. Panel B shows plots by external substitutes defined as the share of employees in the same occupation in the commuting zone relative to the national average (Jäger and Heining, 2019). Panels C and D show regression coefficients of binary indicators for mothers having returned to their pre-birth firm at different points in time. Each point estimate is based on a separate regression. Control variables include at the mothers' level: age dummies, education, citizenship dummy, tenure and pre-birth earnings (both linear and squared). We further include occupation FEs, log firm size, labour market region FEs (141) and the two variables plotted in the panels. 95% confidence level calculated with robust standard errors. Source: IEB, own calculations.

rather than a continuous definition of days until return to trace out the relationship over time after birth without complications due to right-censoring. For our main specification, we use $\ln(\text{internal substitutes})$ and $\ln(\text{external substitutes})$ to estimate the correlation between mother's return behaviour and the availability of internal and external substitutes.²⁷ The vector X_i includes worker and firm characteristics determined ten months prior to childbirth, namely tenure (linear and squared), earnings (linear

²⁷The pattern is robust if we employ indicator variables for the number of internal substitutes as shown in Figure 3.5, see Appendix Figure 3.A.8.

and squared), indicator variables for mothers' occupation, age at birth, education and citizenship, regional labour markets, and (log) firm size.

Panel C in Figure 3.5 presents the δ coefficients from eq. (3.1). The relationship between the availability of internal substitutes and mothers' return to their firm is strongest immediately after the expiration of maternity leave. A one percent increase in the number of internal substitutes is associated with an almost four percentage points lower probability to have returned three months after birth. Over time, the relationship gradually weakens and becomes statistically insignificant. This relationship cannot be explained by differences in important worker, firm, and local labour market characteristics.²⁸ Panel D in Figure 3.5 presents the γ coefficients of eq. (3.1) and reveals no significant link between mothers' return to the firm and the availability of external substitutes in the labour market region in the first three years after childbirth.

Overall, the fewer internal substitutes are available, the shorter mothers go on leave. Additionally, firms make more replacement hirings when women with few internal substitutes become mothers.

3.4.3 Mothers' Return Behaviour with Extended Paid Parental Leave

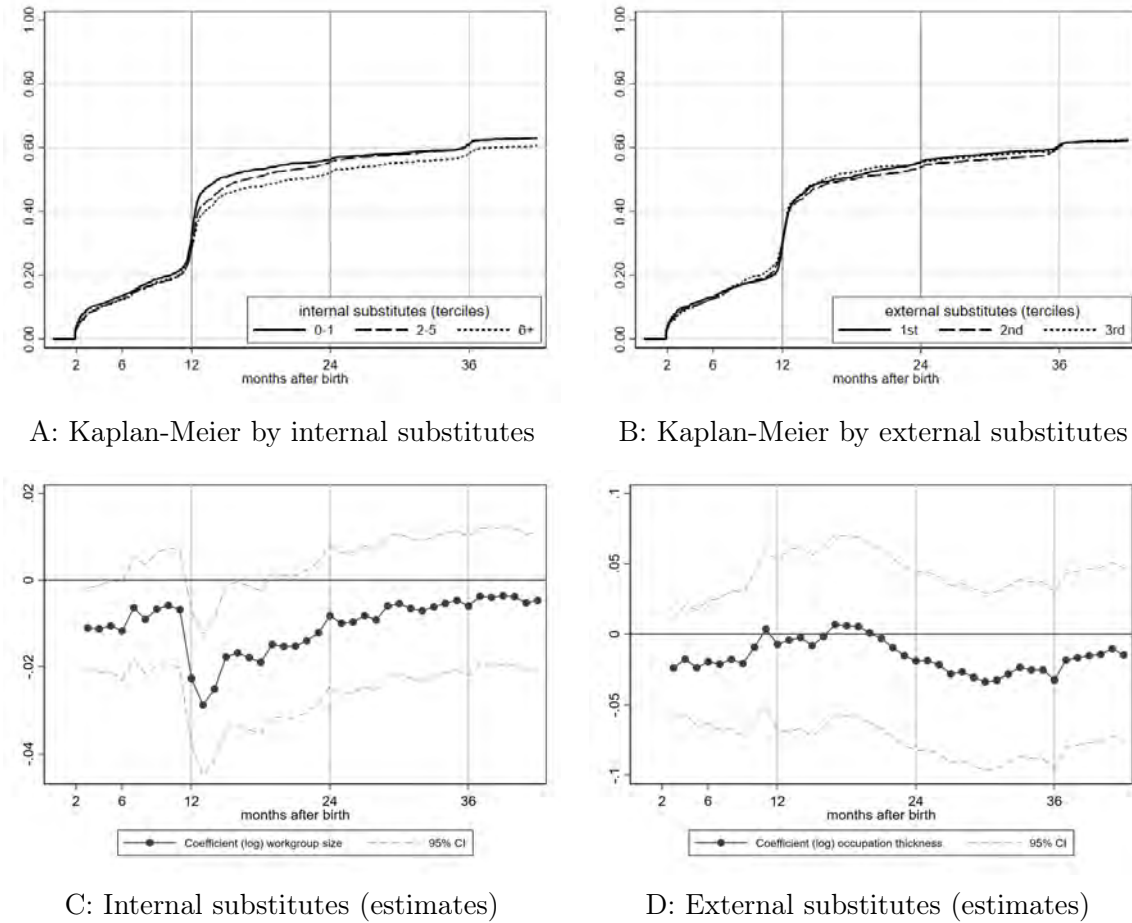
Next, we examine how the introduction of more generous paid parental leave affects the link between mothers' return and the availability of internal and external substitutes. Panels A and B of Figure 3.6 provide the Kaplan–Meier plots for returning to the pre-birth employer distinguishing by the availability of internal and external substitutes for mothers giving birth under the new paid parental leave regime (i.e., between January and June 2007).

The key finding that emerges from Panel A is that mothers with few internal substitutes now return at almost the same rate as mothers with more internal substitutes in the first year after childbirth. Thus, the extension of paid parental leave benefits in the first year after childbirth strongly reduced the previous differences in mothers' return-to-work behaviour by the number of internal substitutes. With respect to the availability of external substitutes, Panel B reveals no systematic differences in mothers' return related to the regional labour market thickness.

In Panels C and D of Figure 3.6, we condition on worker, occupation, and firm characteristics, as well as on labour market regions (see eq. (3.1)). Panel C shows that the link between the availability of internal substitutes and returning to work has weakened substantially and turned statistically insignificant during the first year after childbirth. Once mothers' paid leave expires after 12 months, the return-to-work pattern by the availability of internal substitutes returns to its pre-reform pattern.

²⁸Appendix Table 3.A.5 shows that these relationships are robust to various choices of control variables.

Figure 3.6: Return to pre-birth firm in setting with extended paid leave entitlements



Notes: Sample period is births from January 2007 to June 2007. See Figure 3.5 for other notes.
Source: IEB, own calculations.

Thus, mothers who are harder to replace internally again return more quickly to their previous employer compared to mothers with more internal replacements. Panel D shows that the availability of external substitutes is unrelated to mothers' return when additional control variables are included.

Overall, the introduction of paid parental leave distorted the coordination between employers and mothers on their return in the short-run. This may carry additional costs for firms to bridge the gap in the workforce. In the next sections, we will explore the effects of the parental leave extension on mothers' absences and firms' responses to this shock.

3.5 Effects of Extended Parental Leave Absences on Mothers and Firms

3.5.1 Empirical Strategy

To estimate the effects of the 2007 parental leave reform on mothers and firms, we implement a dynamic difference-in-differences design (similar to Ginja et al., 2020b). We describe the estimation strategy for the effects on mothers and use the same estimation strategy for firms as we observe one birth per firm.

For the first difference, we compare outcomes between mothers giving birth up to six months before and after January 1, 2007. To account for seasonal variations and time trends in outcomes, we take a second difference using mothers giving birth one year earlier, i.e., up to six months before and after January 1, 2006. Moreover, we can use the dynamic evolution of outcomes as an additional difference. This allows us to examine the development of the estimated treatment effects over time and to directly assess any potential pre-treatment differences between treatment and control units.

We estimate the effects of the parental leave reform on mothers' monthly outcomes with the following regression model:

$$\begin{aligned}
 y_{it} = & \sum_{t=-24}^{54} \gamma_t \mathbb{1}(T_t) \times reform_i \times spring_i + \sum_{t=-24}^{54} \delta_t \mathbb{1}(T_t) \times reform_i + \\
 & \sum_{t=-24}^{54} \tau_t \mathbb{1}(T_t) \times spring_i + \sum_{t=-24}^{54} \beta_t \mathbb{1}(T_t) + \epsilon_{it}
 \end{aligned} \tag{3.2}$$

where y is the outcome of mother i at event-time t ; $t = 0$ corresponds to the month of birth.²⁹ The variable $reform_i$ takes the value of 1 if the mother gives birth between July 2006 and June 2007, and 0 otherwise. The variable $spring_i$ indicates whether a birth occurred between January and June of a year. We omit the event time dummy for $t = -10$, so that the coefficients γ_t estimate the treatment effect in each time period t relative to ten months prior to childbirth. We bin the endpoints on either side of the effect window (Schmidheiny and Siegloch, 2020). Standard errors are clustered at the mother-level. For earnings, we use annual earnings and calculate eq. (3.2) in calendar years and compare it to the pre-birth year.

To summarise the effect sizes, we also report the estimates for four discrete time bins. Specifically, we report effects for the *pre-birth period* (24 to 11 months before birth), as

²⁹We omit mother fixed effects from the regression equation, because we use a balanced panel and their inclusion does thus not affect our estimates.

well as *short-term effects* (2 to 14 months after birth) covering the paid parental leave period, *medium-term effects* (15 to 36 months after birth) covering the remaining job protection period, and *longer-term effects* (37 to 54 months after birth). The period from ten months before birth to two months post-birth is the reference period.³⁰ We estimate the following regression:

$$\begin{aligned}
 y_{it} = & \sum_{t=p,s,m,l} \gamma_t^d \times \mathbb{1}(D_t) \times reform_i \times spring_i + \sum_{t=p,s,m,l} \delta_t^d \times \mathbb{1}(D_t) \times reform_i \\
 & + \sum_{t=p,s,m,l} \tau_t^d \times \mathbb{1}(D_t) \times spring_i + \sum_{t=p,s,m,l} \beta_t^d \times \mathbb{1}(T_t) + u_{it}
 \end{aligned} \tag{3.3}$$

where γ_t^d denote the pre-birth (p), short- (s), medium- (m), and longer-term (l) effects.

To estimate the effects on firms, firms replace mothers as the unit of analysis and we define groups as of when the birth occurred in the firm.

Identifying assumptions. To interpret the γ coefficients as the causal effect of the 2007-reform, (i) selection into motherhood must not have changed, (ii) the timing of births around the policy cut-off needs to be as good as random, and (iii) the potential outcomes between treatment and control mothers and firms must follow common trends. The way the reform was passed and empirical findings on its fertility effects support the first assumption: Although the reform was first publicly discussed in May 2006, the final law was only passed in September 2006 (Kluve and Tamm, 2013). All births occurring until June 2007 had been conceived prior to the passing of the reform, such that parents could potentially react with conception only thereafter. In line with this, Raute (2019) observes first fertility responses only from August 2007 onward; as our sample only contains births until June 2007, differential selection into motherhood should not bias our estimates. We substantiate this point empirically in column (6) of Table 3.1, which reports the coefficients from difference-in-differences estimations. The coefficients reveal no systematic differences between the treatment and control groups in mothers' or firms' characteristics.

The second assumption is threatened if mothers shift the timing of births near the reform cut-off by postponing cesarean sections or labour inductions to benefit from the new regulation (Jürges, 2017; Neugart and Ohlsson, 2013; Tamm, 2013). To deal with this concern, we exclude women with expected dates of delivery in the two weeks before and after the cut-off and perform a density test for equally distributed births

³⁰For the annual earnings estimates, the pre-birth period is two years before birth, short-term is the birth year and the following, medium-term is two to three years and longer-term is four years after birth. The calendar year preceding birth serves as the reference period.

near the cut-off following Cattaneo et al. (2018). Our estimates reveal no evidence for any significant birth shifts (see Appendix Figure 3.A.9).

Third, a causal interpretation of our estimates requires that mothers and firms in the treatment and control groups follow a common trend in the evolution of their potential outcomes. As the potential outcomes are not observable, we assess pre-treatment trends in outcomes throughout the analysis and find no meaningful differences.

3.5.2 Worker Absences and Effects on Firms

We begin our quasi-experimental analysis by examining how the reform affected mothers' absences from their workplace after childbirth.³¹ We estimate the reform effect on mothers' labour market outcomes using the dynamic difference-in-differences model outlined in eq. (3.2). Figure 3.7, Panel A, shows how the reform affected mothers' probability to be employed at their pre-birth firm. In the two years before childbirth, we estimate flat pre-trends, which supports our main identification assumption.³² After childbirth, the parental leave reform substantially decreased mothers' probability to work for their pre-birth employers throughout the first year after birth (by a maximum of 20 percentage points ten months after birth).³³ We observe no meaningful medium- or longer-term differences in the probability to work at the same firm up to 54 months after childbirth, see also column (1) of Table 3.2 which summarises the estimates. These findings imply that the reform strongly increased mothers' parental leave absences in the first year after childbirth but had no effect on mothers' long-run absences, e.g., through effects on separations.

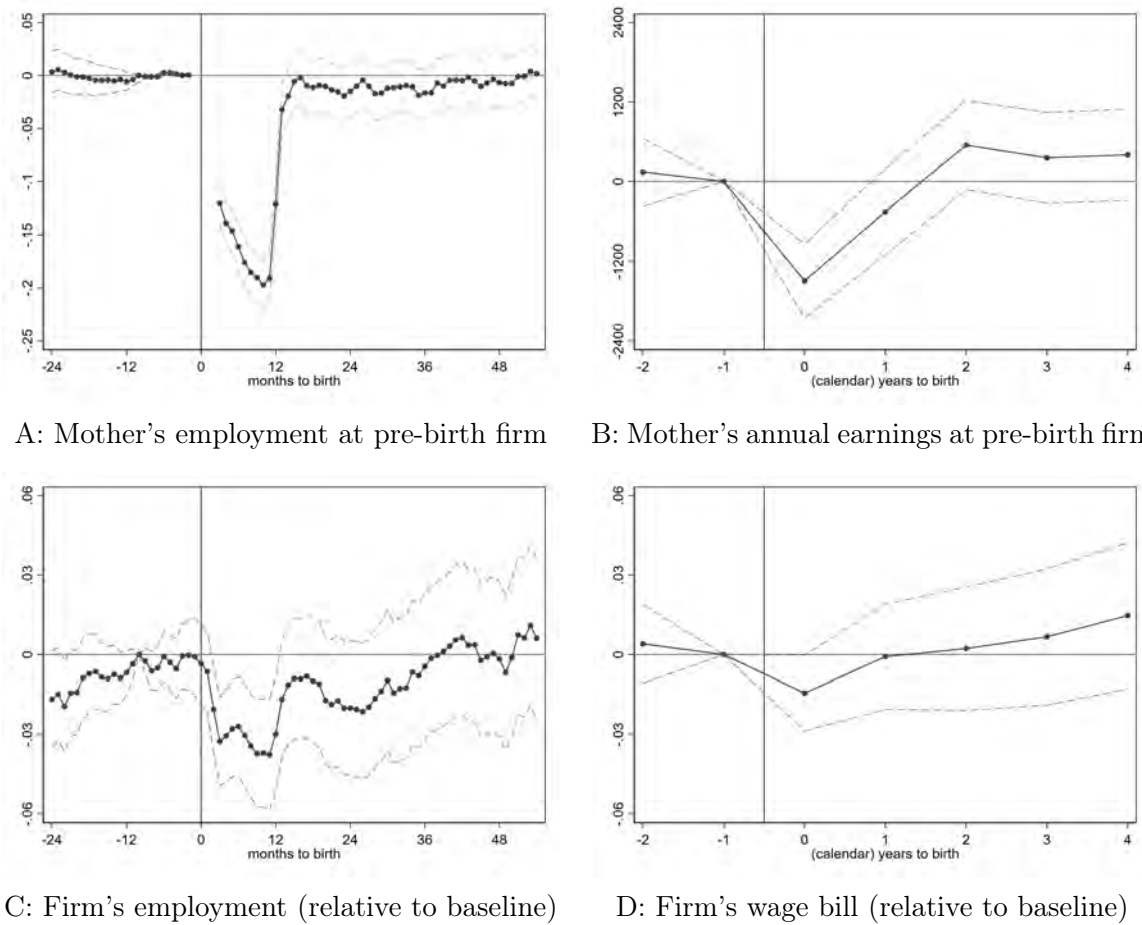
Panel B of Figure 3.7 presents the reform effect estimates on mothers' annual earnings at their pre-birth firms based on eq. (3.2). Treated mothers follow the same earnings trends within their firms prior to childbirth. Consistent with the longer absence after childbirth due to the reform, their earnings drop below those of the control group in the first two years after childbirth. In the following years, the earnings of treated mothers are above the earnings of mothers in the control group, but the difference is small

³¹Several other empirical studies have examined how the reform affected maternal labour market outcomes such as employment and earnings (e.g., see Kluve and Tamm, 2013; Kluve and Schmitz, 2018; Frodermann et al., 2020). For completeness, Appendix Table 3.A.6 reports comparable results for our sample of mothers where we consider employment at all firms. Our focus on the return to the pre-birth employer was also analysed in Kluve and Schmitz (2018), who find that high-earning mothers are more likely to return to their previous employers by 2 percentage points, and they are more likely to hold unlimited contracts.

³²As we condition our sample on mothers working in the same firm at least ten months prior to childbirth (see section 3.3), estimates at time $-10 \leq t \leq 0$ are deterministically close to zero.

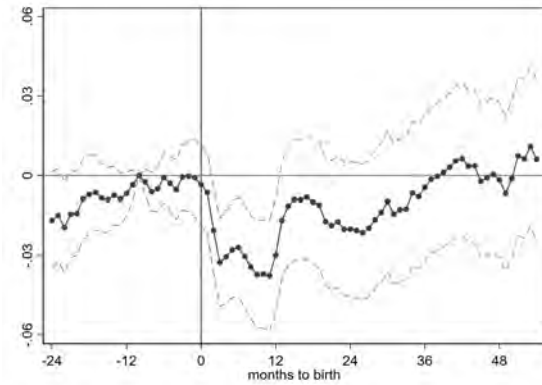
³³Figure 3.1 shows that about 34 percent of mothers giving birth prior to the reform returned to their pre-birth employer within the first six months, this share decreased by 20 percentage points, or 57 percent, for women giving birth after the reform.

Figure 3.7: Event study of parental leave reform effects on mothers' and firms' outcomes

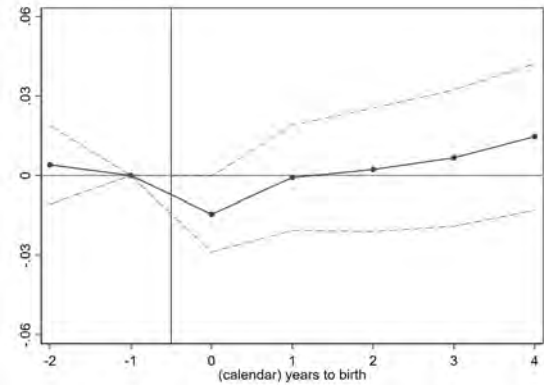


A: Mother's employment at pre-birth firm

B: Mother's annual earnings at pre-birth firm



C: Firm's employment (relative to baseline)



D: Firm's wage bill (relative to baseline)

Notes: The figure plots event study estimates of the 2007 paid parental leave reform in Germany on maternal labour supply and firm outcomes based on eq. (3.2). Dashed lines indicate 95% confidence interval, standard errors clustered at the mother / firm level. Information on earnings in Panels B and D are reported annually; earnings in 2010 euro.

Source: IEB, own calculations.

(around 400 euro) and not statistically significant. Panel B of Table 3.2 presents the corresponding summary estimates in column (1).

Next, we examine how this negative, temporary labour supply shock affects firms' total employment and labour costs. We first examine the gap that mothers' absences create in the firm. In frictionless labour markets, we would expect that firms fully compensate the gap at the extensive margin through deferred separations or increased hirings. Panel C of Figure 3.7 examines employment at the firm and shows that employment is reduced after childbirth in firms exposed to longer maternal absences. Compared to ten months prior to childbirth, the parental leave expansion reduces employment within the first year after childbirth by around three percent in treated firms. The treatment effect turns insignificant 12 months after childbirth and converges

Table 3.2: Summary event study estimates - mothers

	All	Internal substitutes		
		0-1	2-5	6+
<u>Panel A: Employed at pre-birth firm</u>				
Pre-period	-0.001 (0.006)	-0.003 (0.011)	0.015 (0.010)	-0.018 (0.011)
Short-term effect	-0.132*** (0.009)	-0.148*** (0.016)	-0.132*** (0.015)	-0.109*** (0.017)
Medium-term effect	-0.011 (0.011)	-0.006 (0.019)	-0.007 (0.018)	-0.022 (0.021)
Longer-term effect	-0.002 (0.011)	-0.006 (0.019)	0.005 (0.019)	-0.007 (0.021)
Mothers	23,679	8,624	8,504	6,551
Observations	2,415,258	879,648	867,408	668,202
<u>Panel B: Annual earnings in calendar year at pre-birth firm</u>				
Pre-period	140.873 (261.608)	73.032 (449.575)	556.394 (419.608)	-279.480 (495.300)
Short-term effect	-971.978*** (287.967)	-1038.139** (479.027)	-966.096** (465.193)	-923.822 (565.502)
Medium-term effect	453.893 (334.729)	687.969 (558.677)	273.447 (540.398)	345.682 (656.514)
Longer-term effect	402.304 (351.344)	601.613 (587.651)	21.743 (564.223)	602.997 (691.606)
Mothers	23,679	8,624	8,504	6,551
Observations	189,432	68,992	68,032	52,408

Notes: Table summarises event study estimates for the main outcomes of mothers in discrete time periods based on eq. (3.3). Estimates in Panel A are based on monthly information. Pre-birth is from 28 to 11 months pre-birth, the period from ten months pre- to one months post-birth is the omitted period. Short-, medium- and longer-term refer to 2-14, 15-36 and 37-58 months post-birth, respectively. For the annual estimation in Panels C and D, pre-birth is two calendar years before birth, we omit the year before and short-, medium- and longer-term refer to 0-1, 2-3 and 4 years after birth. Standard errors clustered at the mother level in parentheses. Significance levels: * < 10% ** < 5% *** < 1%.
Source: IEB, own calculations.

to zero within three years after childbirth, that is after the expiry of the job protection period. Firms' total labour costs are not statistically significantly affected (Figure 3.7, Panel D), though the negative estimate in the year of childbirth suggests that mothers are not fully replaced through intensive margin adjustments. Column (1) of Table 3.3 provides corresponding short-, medium- and longer-term estimates.

Next, we examine treatment effect heterogeneities and analyse whether the effects on firms differ by the availability of internal substitutes for the mother on leave. Panel A of Figure 3.8 shows that the reform reduces total employment at the firm when at most one internal substitute is available. Table 3.3 shows that employment at these firm drops by about 3.4 percent in the first 14 months after childbirth and their wage bill drops by about 1.6 percent in the year of childbirth and the following year. The

Table 3.3: Summary event study estimates - firm

	All	Internal substitutes		
		0-1	2-5	6+
Panel A: Firm's relative employment				
Pre-period	-0.008 (0.007)	-0.012 (0.014)	-0.006 (0.010)	-0.003 (0.010)
Short-term effect	-0.026*** (0.007)	-0.034** (0.014)	-0.029*** (0.011)	-0.010 (0.011)
Medium-term effect	-0.011 (0.011)	-0.021 (0.021)	-0.018 (0.017)	0.011 (0.018)
Longer-term effect	0.007 (0.013)	-0.011 (0.025)	0.008 (0.021)	0.030 (0.023)
Firms	23,679	8,624	8,504	6,551
Observations	2,415,258	879,648	867,408	668,202
Panel B: Firm's relative annual wage bill				
Pre-period	0.004 (0.008)	-0.004 (0.015)	0.014 (0.012)	0.004 (0.012)
Short-term effect	-0.008 (0.008)	-0.016 (0.016)	-0.006 (0.012)	0.002 (0.013)
Medium-term effect	0.004 (0.012)	-0.011 (0.022)	0.007 (0.019)	0.021 (0.021)
Longer-term effect	0.015 (0.014)	-0.023 (0.025)	0.031 (0.023)	0.043* (0.025)
Firms	23,679	8,624	8,504	6,551
Observations	189,432	68,992	68,032	52,408

Notes: The table summarises event study estimates for the main outcomes of at the firm level in discrete time periods based on eq. (3.3). See Table 3.2 for other notes. Standard errors clustered at the firm level in parentheses. Significance levels: * < 10% ** < 5% *** < 1%.

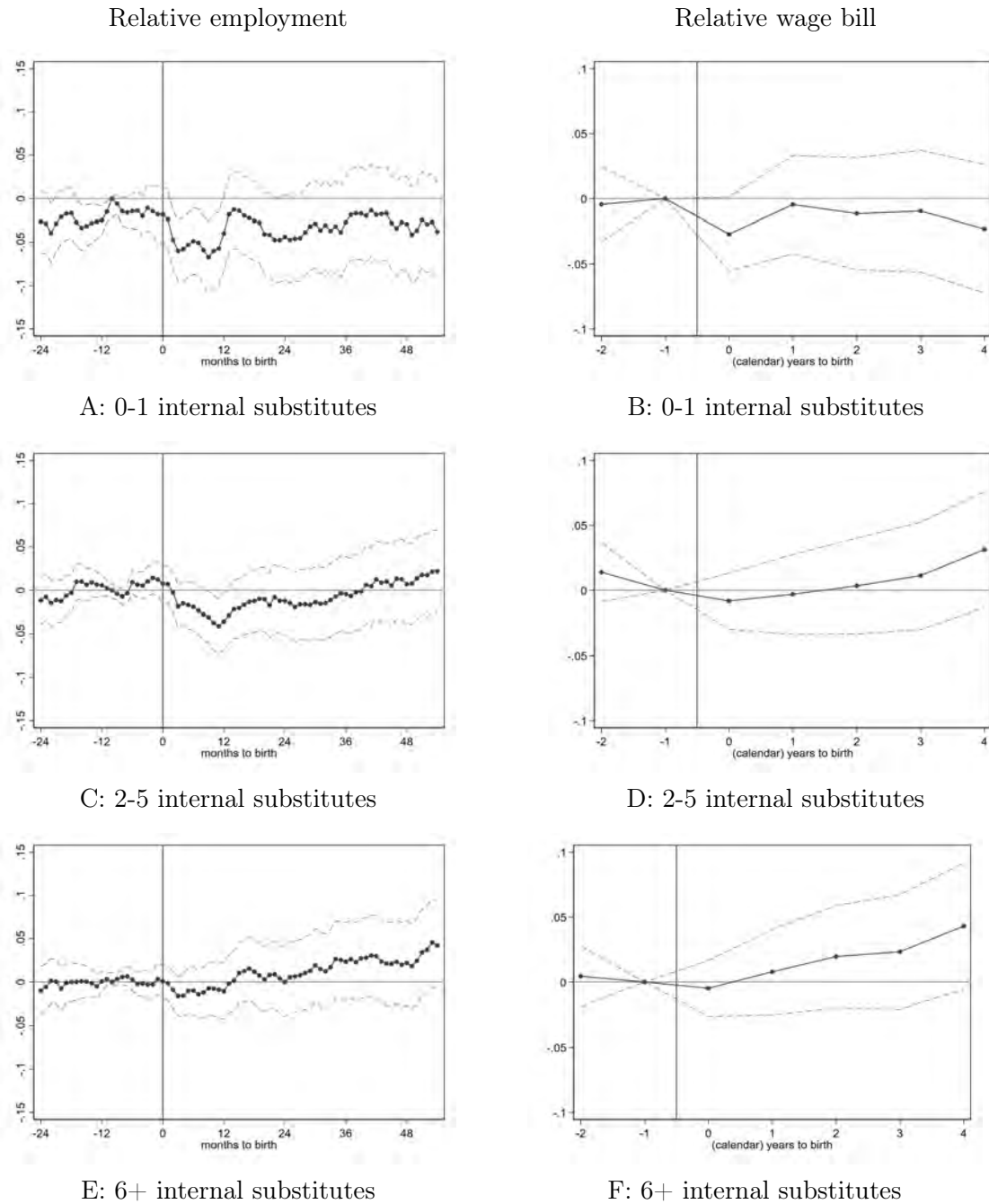
Source: IEB, own calculations.

employment gap reduces over time and turns statistically insignificant in the medium- and longer-term.

With 2-5 internal substitutes (Panel C), firms' employment also drops by three per cent when the mother goes on extended leave. This drop is similar to the effect on firms with fewer internal substitutes. We find no effect on the wage bill for this group (Panel D). However, in firms with 6 or more internal substitutes for the mother on leave, we do not observe any drop in employment or the wage bill (Panels E and F). Though the longer run estimates are less precise and not statistically significant, the point estimates suggest that firms with six or more internal substitutes may have benefited from the parental leave reform.

Although the differences between the groups are not statistically significant, the point estimates indicate substantial effect heterogeneity depending on the number of

Figure 3.8: Effect heterogeneity - event study of parental leave reform effects on firm outcomes



Notes: The figure plots event study estimates of the 2007 paid parental leave reform in Germany on firm outcomes based on eq. (3.2) (with 95% confidence interval), separately by the size of the workgroup in which the birth took place. The baseline month for employment effects is ten months prior to childbirth, and for the wage bill one calendar year prior to childbirth. Wage bill in 2010 euro.
Source: IEB, own calculations.

internal substitutes for the mother. Overall, the point estimates support the conclusion that firms that have few internal substitutes available cannot fully close the labour

shortage gap caused by longer parental leave absences. In our final section, we investigate whether firms internalise these costs through statistical discrimination.

3.6 Effects of Extended Parental Leave on Hiring Decisions

We have shown that the paid parental leave expansion created a short-term gap in firms' employment when few internal substitutes are available for the mother-on-leave. We now analyse whether the paid parental leave expansion also affected the hiring composition of firms. We hypothesise that profit-maximising firms anticipate and internalise the potential costs of longer birth-related absences by younger female workers and in turn reduce the hiring of younger women.³⁴ Thus, finding that firms after the reform hire fewer women of childbearing age for workgroups with few internal substitutes would indicate statistical discrimination.

For this analysis, we modify the empirical approach and now study the hiring behaviour by firms in our sample into workgroups that are *not* directly affected by a birth.³⁵ The advantage of examining these workgroups is that they were not directly impacted by the childbirth occurring in another occupation in the same firm. Thus, their hirings should not (or to a lesser degree) be distorted by having to replace a mother going-on-leave. We distinguish between four mutually exclusive groups of hirings: By sex and by age, specifically above and below age 38. We consider women below age 38 as potential mothers as most women have completed their fertility by then.³⁶

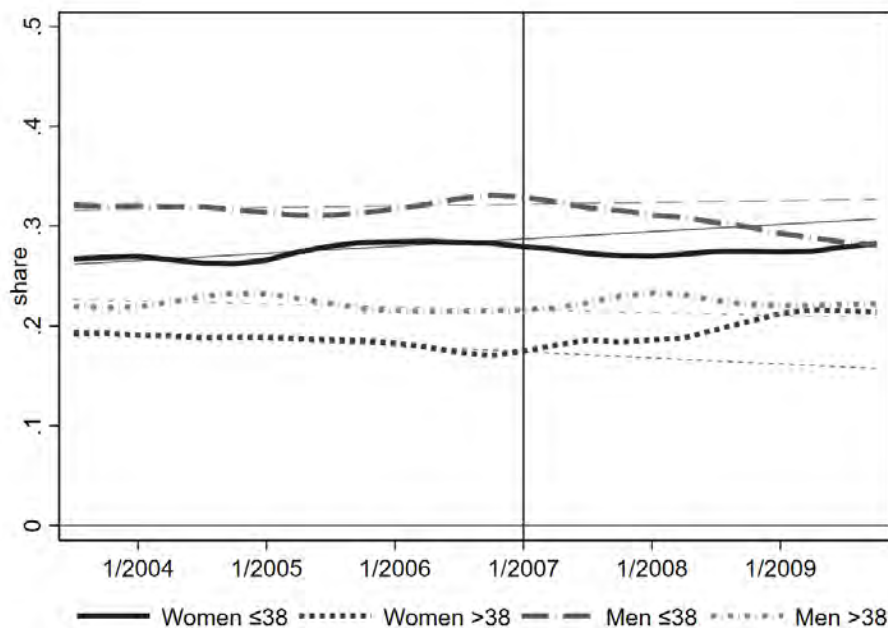
To illustrate the intuition of our approach, which follows Dobkin et al. (2018), we plot the composition of all hiring events for these groups by calendar time in Figure 3.9. For this figure, we net out calendar month effects and estimate a linear trend over the pre-reform period from July 2003 to December 2006. We extrapolate this trend separately for each group over the entire sample period as the counterfactual trend in hirings.³⁷ To estimate the effect of the reform on the composition of hirings, we then estimate the monthly deviations from the pre-reform time trends.

Overall, Figure 3.9 shows that aggregate hiring shares of all demographic groups follow fairly linear trends in the pre-reform period. In the absence of the parental leave reform, we would expect hiring shares to continue along these paths. However, the aggregate figure already shows that the hiring patterns start to diverge from their

³⁴Statistical discrimination may not only affect actual or future mothers. Fernández-Kranz and Rodríguez-Planas (2021) find that a right for part-time work for young mothers had negative hiring effects on young childless women as well.

³⁵We impose the condition that firms had at least one worker in a specific occupation on June 30, 2006, i.e., half a year before the parental leave extension. The condition that workgroups existed at a uniform reference date before the reform ensures that time-variant effects do not impact workgroups

Figure 3.9: Composition of hirings



Notes: Figure shows the residualised (calendar month effects are partialled out) demographic composition of all hirings at a quarterly level. Sample consists of all workgroups in sample firms, in which no birth is observed. Trend lines are calculated based on the period before the parental leave expansion in January 2007.

Source: IEB, own calculations.

pre-reform trend with the introduction of the parental leave reform. In particular, we observe that the hiring shares of young workers, both male and female, decline after the reform.

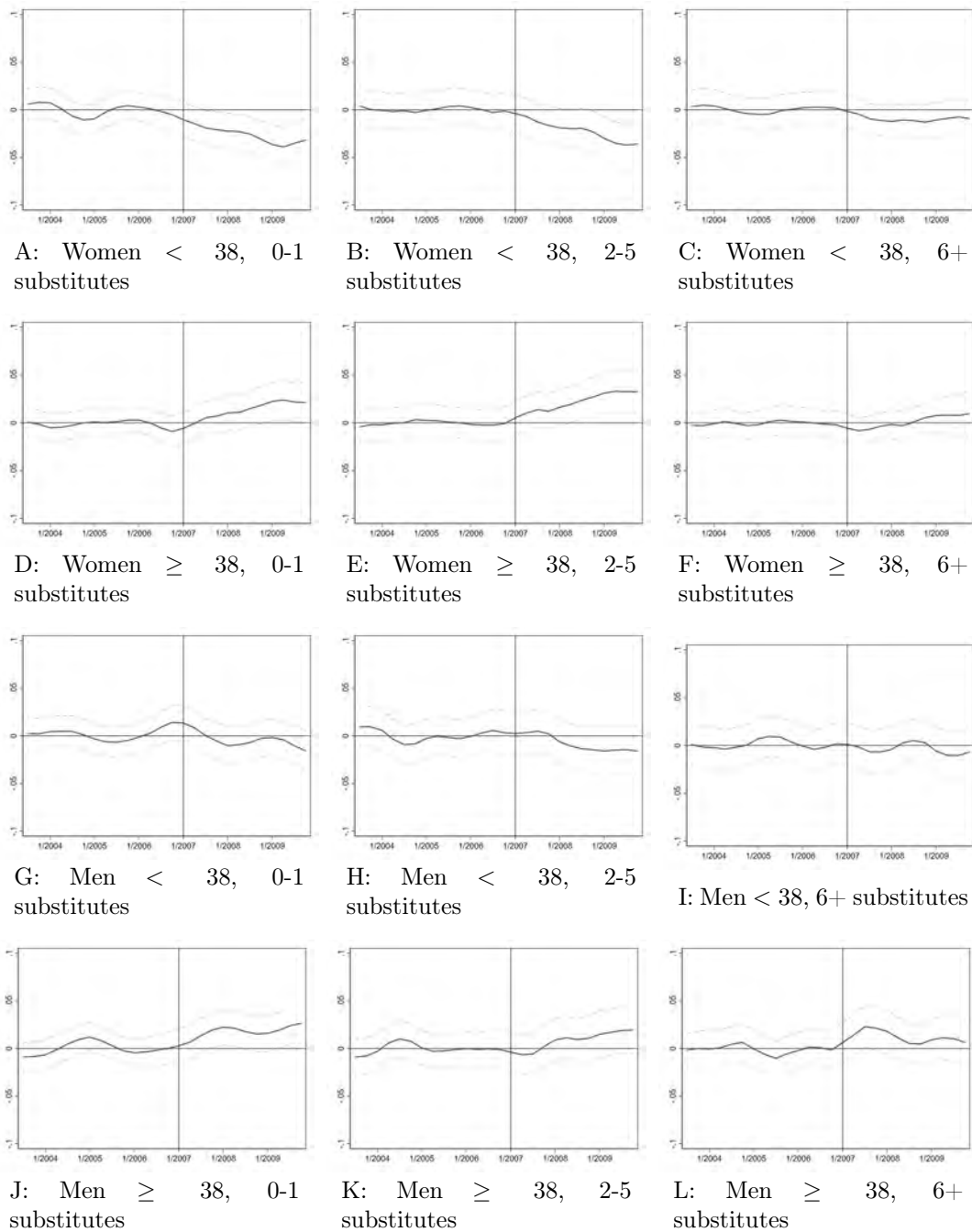
To disentangle whether these patterns depend on the availability of internal substitutes, Figure 3.10 plots the monthly deviations from the pre-reform time trend separately for each demographic group and by the availability of internal substitutes. In these estimations, we control for occupation fixed effects given that men and women work in different occupations. We estimate flat pre-trends across all panels, which supports our identification strategy. Starting with Panel A, we find that the share of young women among new hirings declines in workgroups with fewer internal substitutes (Panels A and B). Conversely, we estimate an increase in the share of hirings for older women in firms with fewer internal substitutes (Panels D and E). These pattern

differently if the conditions is set on an existence at different time periods. In total, from July 2003 to December 2009 we observe 388,132 hiring events in 78,006 workgroups in 18,799 firms.

³⁶According to the Federal Statistical Office, only 12 percent of births are from mothers above 38, most of which are higher order births. Following Müller and Strauch (2017), we use the same restriction of 38 years to identify first births in the administrative data.

³⁷We weigh by the inverse of the number of hirings per workgroup, to give each group equal weight in this analysis (analogously to earlier analyses).

Figure 3.10: Trend deviations in hiring composition by demographic group and internal substitutes



Notes: Figure shows quarterly deviations from a linear trend estimated over the pre-reform period (up to the 4th quarter of 2006) and extrapolated over the entire period. We include occupation FEs in the estimation for this figure.
Source: IEB, own calculations.

are evidence for statistical discrimination. In contrast, we estimate no changes in the hiring shares of young or older women in firms with many internal substitutes (Panels

C and F) which could reflect that these firms are better able to compensate the labour shortage. We do not find any effects for young men (Panels G-I), but observe small increases from their pre-reform time trend for older men (Panels J-L).

3.7 Conclusion

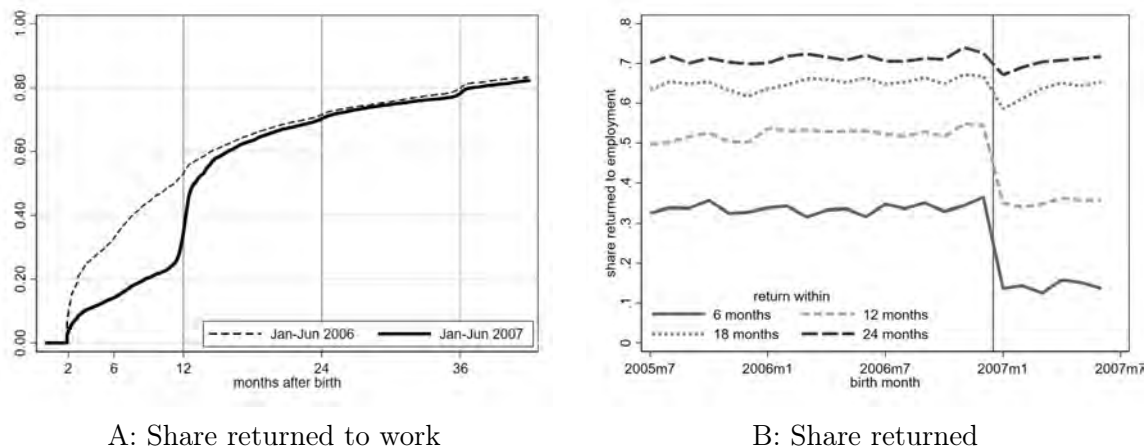
This paper takes a firm-side perspective on motherhood and parental leave yielding five key insights. First, firms hire more substitutes for the mother on leave when no or few internal substitutes are available (0.3 replacement hirings per mother) compared to firms where more internal substitutes are available. Second, mothers with few internal substitutes take shorter leave than mothers with more internal substitutes when parental leave benefits are limited. Third, this pattern is almost erased by a paid parental leave reform granting more generous benefits in the first year after childbirth. Fourth, the reform delays the return of mothers to their pre-birth employers and this reduces firms' employment in the short-term if they have few internal substitutes, but not in the longer-term. Fifth, firms responded to the reform by reducing the share of young women among hirings into occupations where few internal substitutes are available.

Taken together, our study shows that motherhood and parental leave policies burden firms in the short-term when few internal substitutes for the mother are available. Our analysis exploring statistical discrimination carries the important policy implication that parental leave policies, meant to improve the well-being of families, may backfire on potential mothers. To alleviate the effects of motherhood and parental leave on firms and reduce the scope for statistical discrimination, policy makers could additionally compensate firms for birth-related worker absences.

3.A Appendix

3.A.1 Figures

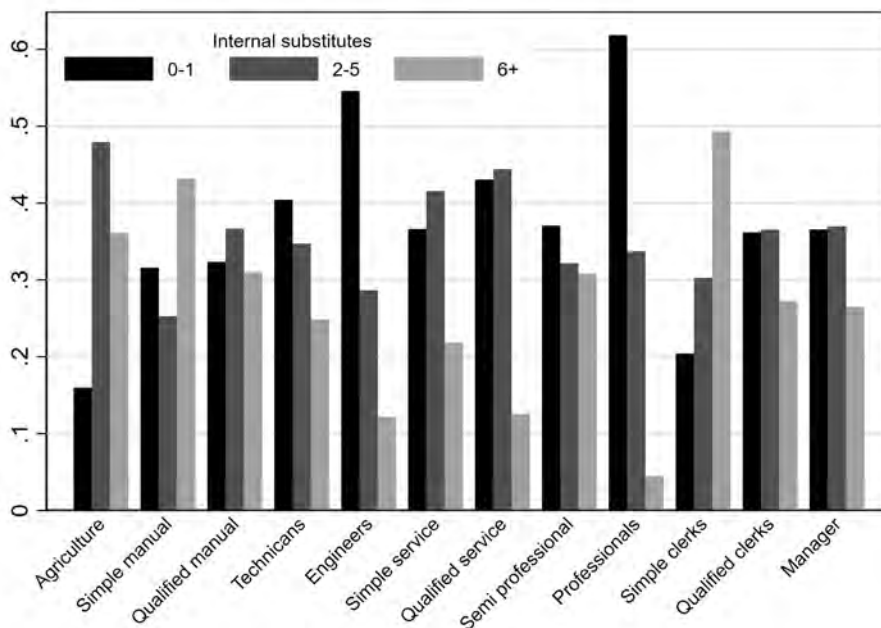
Figure 3.A.1: Return to employment before and after the 2007 parental leave reform



Notes: Panel A plots the share of mothers that have returned to employment by month t after childbirth. The dotted line indicates mothers giving birth between January and June 2006 (pre-reform), the solid line indicates mothers giving birth between January and June 2007 (post-reform). Panel B shows the share of mothers who have returned to employment at discrete points in time.

Source: IEB, own calculations.

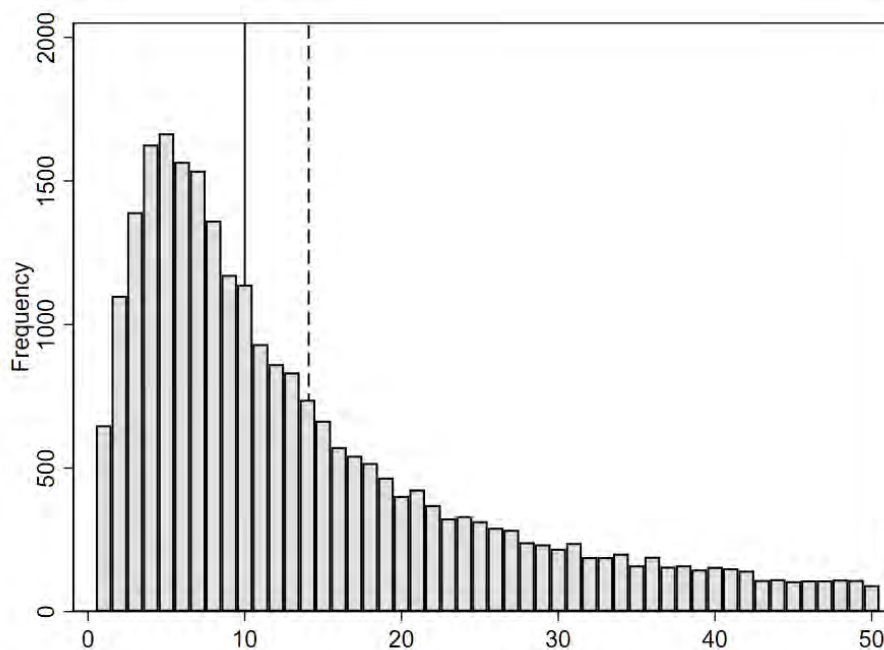
Figure 3.A.2: Occupations by internal substitutes



Notes: Figure shows mothers' occupations by the number of internal substitutes in her firm.

Source: IEB, own calculations.

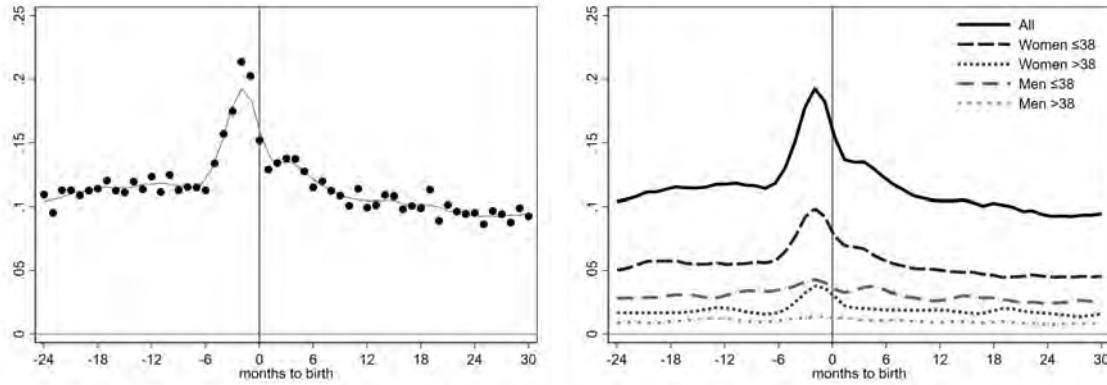
Figure 3.A.3: Size of firms



Notes: Figure shows the distribution of pre-birth firm sizes. The solid line indicates the median firm size, the dashed line the mean size.

Source: IEB, own calculations.

Figure 3.A.4: Hirings around childbirth - 2007 births

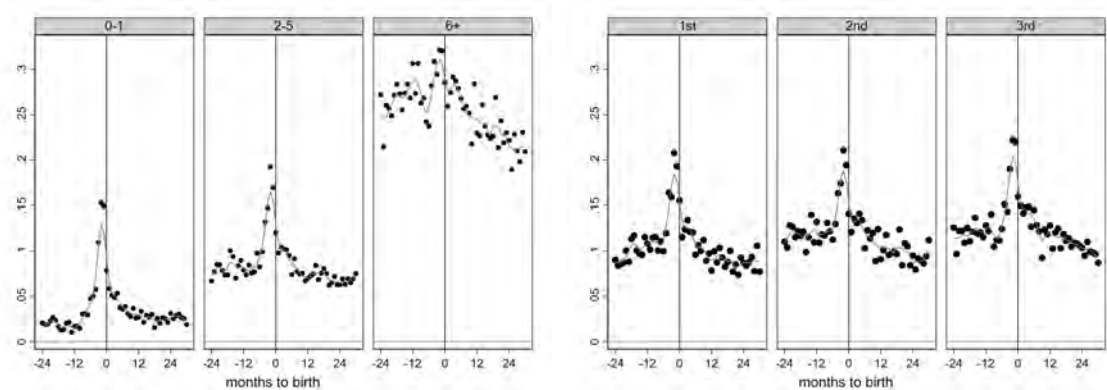


A: Replacement hiring around childbirth

B: Composition of replacement hirings

Notes: Figure shows hiring in the same occupation of mothers around childbirth. Mothers returning to their pre-birth workgroup are not counted as hirings. Numbers of hirings are cleansed of calendar month effects. Sample period is births from January 2007 to June 2007. Source: IEB, own calculations.

Figure 3.A.5: Hirings by availability of substitutes - 2007 births

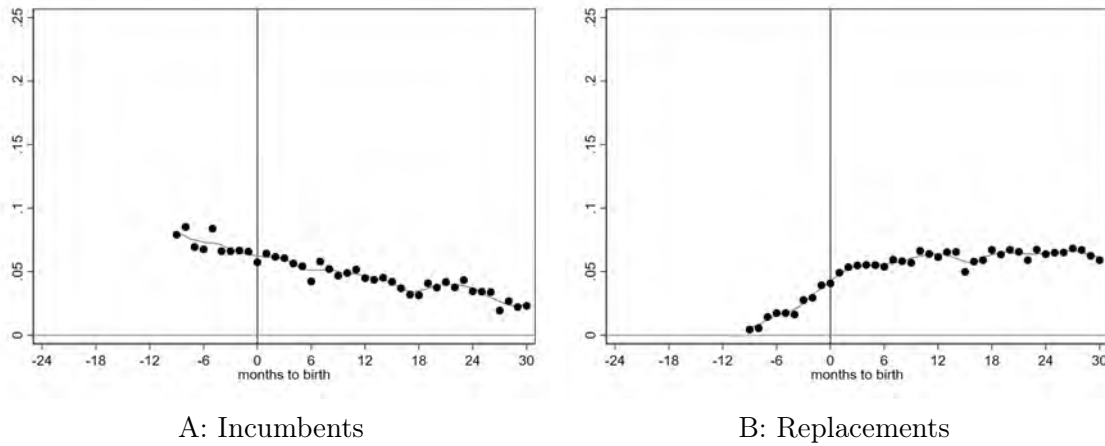


A: Number of internal substitutes (terciles)

B: External substitutes (thickness terciles)

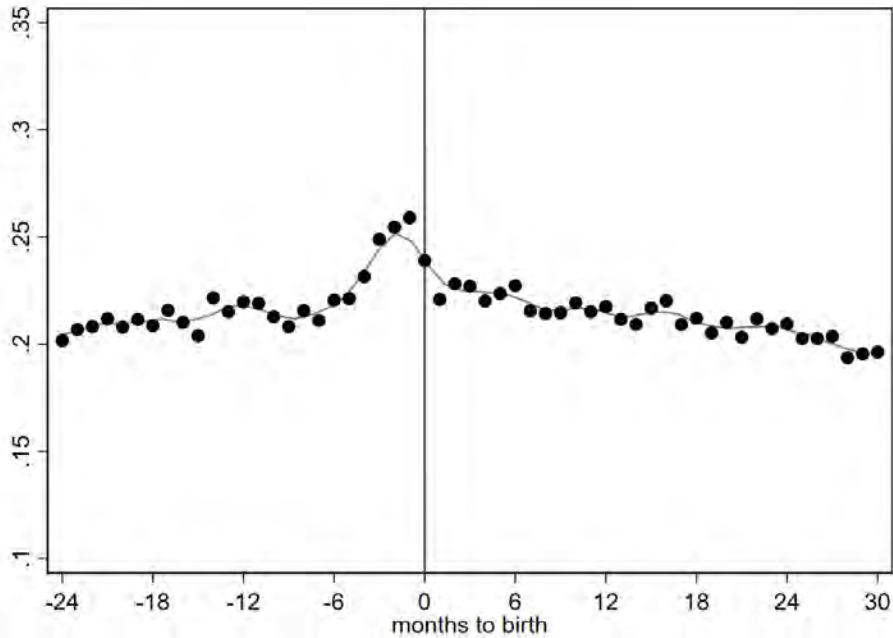
Notes: Figure shows residualised hiring by availability of internal and external substitutes. Internal substitutes are defined as the number of coworkers in the same occupation ten months pre-birth. External substitutes distinguished by terciles of labour market thickness of the occupation. Sample period is births from January 2007 to June 2007. Source: IEB, own calculations.

Figure 3.A.6: Separations by incumbents and replacements - 2007 births



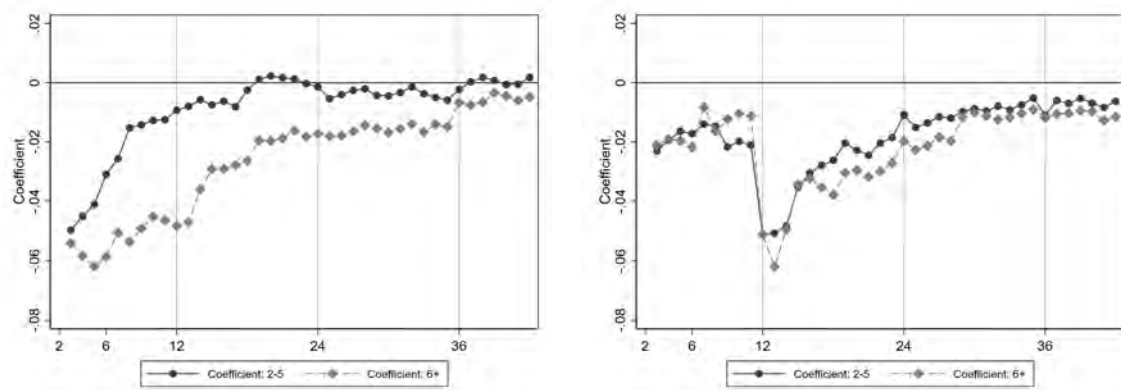
Notes: Figure shows residualised separations of workers in the same occupation as mothers. Panel A shows separation events for incumbents, defines as coworkers ten months before birth. Replacements in Panel B are those hired afterwards. Sample period is births from January 2007 to June 2007. Source: IEB, own calculations.

Figure 3.A.7: Hirings around childbirth in other occupations



Notes: Figures shows hiring in mothers' firms in all other occupations. Number of hirings are cleansed of calendar month effects. Sample period is births from January 2006 to June 2006. Source: IEB, own calculations.

Figure 3.A.8: Return to pre-birth firm and availability of internal substitutes - discrete categorisation

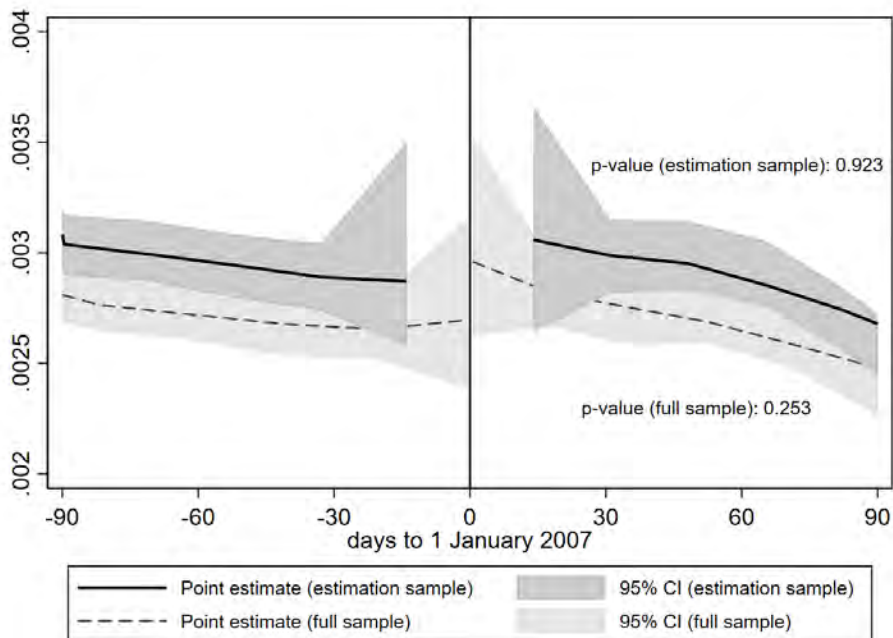


A: Low paid leave (Jan-Jun 2006)

B: Extended paid leave (Jan-Jun 2007)

Notes: Figure shows regression coefficients of binary indicators for mothers to have returned to their pre-birth employer at different points in time. In contrast to panel C of Figure 3.5 this figure shows estimates for the availability of internal substitutes using discrete categories of 2-5 and more than 6 (with 0-1 internal substitutes being the baseline) rather than the continuous definition in Figure 3.5. See Figure 3.5 for other notes.
Source: IEB, own calculations.

Figure 3.A.9: Density test of births



Notes: Figure plots the density of births around the introduction of the paid parental leave on 1 January 2007. The estimation sample excludes the two weeks around 1 January to avoid misassignment of births (see text). Density estimation based on Cattaneo et al. (2018).

3.A.2 Tables

Table 3.A.1: Comparison of mothers in analysis sample with excluded observations

	Analysis sample (1)	Dropped observations (2)
Age at birth	29.96 (4.00)	28.62 (4.89)
German citizen	0.96 (0.20)	0.90 (0.30)
High education	0.39 (0.49)	0.31 (0.46)
Monthly earnings, ten months pre-birth	2,664.23 (786.10)	2,125.26 (1367.14)
At same firm, ten months pre-birth	1.00 (0.00)	0.88 (0.32)
Tenure at current firm in years	4.65 (3.78)	3.80 (3.80)
Full-time employed	0.94 (0.24)	0.82 (0.39)
Non-routine job	0.38 (0.48)	0.37 (0.48)
Return to employment within one year	0.48 (0.50)	0.42 (0.49)
Return to employment within three years	0.79 (0.41)	0.76 (0.43)
Return to pre-birth firm within one year	0.41 (0.49)	0.34 (0.47)
Return to pre-birth firm within three years	0.62 (0.49)	0.58 (0.49)
Observations	23,679	197,995

Notes: Table shows mean values of individual mother characteristics and their pre-birth firm characteristics. Column (1) contains the analysis sample, column (2) consist of all first-time mothers in the analysis period (July 2005 - June 2007) identified in the data that were excluded. The sample restrictions leading to the exclusion are; employed at pre-birth firm ten months before birth, monthly earnings ≥ 1704 euro, one birth in firm in sample period, no public sector and no firms with more than 50 employees pre-birth. Dropped observations exclude mothers from public sector establishments. Source: IEB, own calculations.

Table 3.A.2: Comparison of firms in analysis sample with excluded observations

	Analysis sample (1)	Dropped observations (2)
<i>Characteristics</i>		
Location in West Germany	0.90 (0.31)	0.82 (0.38)
Number of employees	14.53 (12.38)	89.98 (403.89)
Share of female employees	0.61 (0.29)	0.62 (0.29)
Average age of full-time employees	38.57 (5.94)	37.30 (6.15)
Median monthly earnings of full-time employees	2,563.39 (1007.98)	2,174.32 (1092.55)
<i>Sector</i>		
Agriculture, fishing and mining	0.01 (0.11)	0.02 (0.12)
Manufacturing	0.12 (0.33)	0.17 (0.38)
Electricity, gas, water	0.00 (0.06)	0.00 (0.07)
Construction	0.04 (0.20)	0.02 (0.16)
Wholesale and retail	0.33 (0.47)	0.29 (0.45)
Hotels and restaurants	0.02 (0.14)	0.09 (0.29)
Transport, storage, communication	0.05 (0.22)	0.04 (0.21)
Financial intermediation	0.07 (0.25)	0.04 (0.20)
Real estate, renting, business activities	0.30 (0.46)	0.20 (0.40)
Observations	23,679	109,591

Notes: Table compares firm characteristics of the analysis sample with those from observations that were dropped, see Appendix Table 3.A.1.

Source: IEB, own calculations.

Table 3.A.3: Summary statistics by internal substitutes

	Number of internal substitutes			
	0-1	2-5	6+	All
<i>Individual characteristics</i>				
Age in years	30.22 (3.95)	29.84 (3.97)	29.78 (4.08)	29.96 (4.00)
German citizenship	0.96 (0.19)	0.96 (0.20)	0.95 (0.21)	0.96 (0.20)
High education	0.44 (0.50)	0.37 (0.48)	0.35 (0.48)	0.39 (0.49)
Annual earnings in year before birth	30,166.98 (10597.15)	30,152.96 (9888.64)	31,533.23 (10499.92)	30,539.93 (10338.66)
Tenure at current firm in years	4.247 (3.573)	4.814 (3.831)	4.961 (3.918)	4.648 (3.776)
Full-time employed	0.95 (0.22)	0.94 (0.23)	0.93 (0.26)	0.94 (0.24)
<i>Pre-birth firm characteristics</i>				
Location in West Germany	0.88 (0.32)	0.90 (0.30)	0.91 (0.29)	0.90 (0.31)
Firm size	9.56 (9.59)	12.64 (10.15)	21.84 (11.44)	14.06 (11.47)
Workgroup size	1.45 (0.50)	4.17 (1.08)	13.40 (7.53)	5.73 (6.33)
Share of female employees	0.646 (0.312)	0.638 (0.285)	0.602 (0.254)	0.631 (0.288)
<i>Sector</i>				
Agriculture, fishing and mining	0.01 (0.11)	0.01 (0.11)	0.01 (0.11)	0.01 (0.11)
Manufacturing	0.12 (0.33)	0.13 (0.34)	0.12 (0.32)	0.12 (0.33)
Electricity, gas, water	0.00 (0.05)	0.00 (0.05)	0.01 (0.08)	0.00 (0.06)
Construction	0.06 (0.24)	0.04 (0.19)	0.02 (0.13)	0.04 (0.20)
Wholesale and retail	0.33 (0.47)	0.34 (0.47)	0.32 (0.47)	0.33 (0.47)
Hotels and restaurants	0.02 (0.13)	0.02 (0.14)	0.02 (0.14)	0.02 (0.14)
Transport, storage, communication	0.04 (0.19)	0.05 (0.22)	0.07 (0.26)	0.05 (0.22)
Financial intermediation	0.03 (0.17)	0.06 (0.24)	0.12 (0.32)	0.07 (0.25)
Real estate, renting, business activities	0.32 (0.47)	0.31 (0.46)	0.28 (0.45)	0.30 (0.46)
Observations	8,624	8,504	6,551	23,679

Notes: Table shows pre-determined characteristics of the mother and her pre-birth firm by availability of internal substitutes.

Source: IEB, own calculations.

Table 3.A.4: Balancing by internal substitutes
(DD coefficients)

	All	Number of internal substitutes		
		0-1	2-5	6+
<i>Individual characteristics</i>				
Age in years	-0.079 (0.104)	-0.194 (0.170)	-0.092 (0.172)	0.083 (0.202)
German citizenship	0.001 (0.005)	-0.001 (0.008)	-0.002 (0.009)	0.006 (0.011)
High education	0.003 (0.013)	0.007 (0.021)	-0.004 (0.021)	0.004 (0.024)
Annual earnings in year before birth	-108.195 (269.211)	-675.120 (456.486)	356.566 (428.923)	92.954 (521.423)
Tenure at current firm in years	-0.002 (0.098)	0.194 (0.154)	-0.205 (0.166)	0.030 (0.194)
Full-time employed	0.000 (0.006)	0.009 (0.010)	-0.022** (0.010)	0.017 (0.013)
<i>Pre-birth firm characteristics</i>				
Location in West Germany	0.006 (0.008)	-0.000 (0.014)	0.006 (0.013)	0.015 (0.015)
Firm size	0.172 (0.298)	-0.118 (0.414)	0.268 (0.441)	0.584 (0.566)
Workgroup size	0.040 (0.165)	-0.022 (0.021)	0.017 (0.047)	0.310 (0.373)
Share of female employees	0.002 (0.007)	0.011 (0.013)	-0.003 (0.012)	-0.006 (0.013)
<i>Sector</i>				
Agriculture, fishing and mining	0.001 (0.003)	0.003 (0.005)	0.003 (0.005)	-0.005 (0.006)
Manufacturing	0.006 (0.009)	0.005 (0.014)	0.006 (0.015)	0.007 (0.016)
Electricity, gas, water	-0.001 (0.002)	0.001 (0.002)	0.001 (0.002)	-0.006 (0.004)
Construction	-0.013** (0.005)	-0.025** (0.010)	-0.010 (0.008)	0.001 (0.006)
Wholesale and retail	-0.017 (0.012)	0.011 (0.020)	-0.035* (0.021)	-0.030 (0.023)
Hotels and restaurants	0.003 (0.004)	0.005 (0.006)	-0.002 (0.006)	0.007 (0.007)
Transport, storage, communication	-0.011* (0.006)	-0.009 (0.008)	-0.017* (0.009)	-0.005 (0.013)
Financial intermediation	-0.008 (0.006)	0.002 (0.007)	-0.010 (0.010)	-0.018 (0.016)
Real estate, renting, business activities	0.017 (0.012)	0.001 (0.020)	0.034* (0.020)	0.016 (0.022)
Observations	23,679	8,624	8,504	6,551

Notes: The table shows difference-in-differences coefficients for pre-determined characteristics by size of the workgroup. Column (1) corresponds to column (6) of Table 3.1. Robust standard errors in parentheses. * < 10% ** < 5% *** < 1%.

Source: IEB, own calculations.

Table 3.A.6: Summary event study estimates: Mothers - employment at any firm

	All	Internal substitutes		
		0-1	2-5	6+
Panel A: Employed				
Pre-period	-0.002 (0.003)	-0.004 (0.006)	0.003 (0.005)	-0.004 (0.005)
Short-term effect	-0.157*** (0.009)	-0.175*** (0.016)	-0.160*** (0.015)	-0.129*** (0.017)
Medium-term effect	-0.005 (0.010)	0.007 (0.017)	-0.006 (0.017)	-0.019 (0.020)
Longer-term effect	0.011 (0.010)	0.035** (0.017)	-0.014 (0.017)	0.011 (0.020)
Mothers	23,679	8,624	8,504	6,551
Observations	2,415,258	879,648	867,408	668,202
Panel B: Annual earnings in calendar year				
Pre-period	-338.123* (172.954)	-548.526* (307.238)	-107.958 (277.925)	-331.242 (309.212)
Short-term effect	-1349.159*** (266.316)	-1329.301*** (448.312)	-1303.258*** (426.642)	-1476.589*** (519.195)
Medium-term effect	709.847** (327.543)	1360.123** (555.344)	290.144 (525.388)	353.258 (632.918)
Longer-term effect	687.465** (347.236)	1877.771*** (591.582)	-243.651 (556.142)	318.223 (668.018)
Mothers	23,679	8,624	8,504	6,551
Observations	189,432	68,992	68,032	52,408

Notes: Table summarises event study estimates for the main outcomes of mothers in discrete time periods based on eq. (3.3). Estimates in Panel A re based on monthly information. Pre-birth is from 28 to 11 months pre-birth, the period from ten months pre- to one months post-birth is the omitted period. Short-, medium- and longer-term refer to 2-14, 15-36 and 37-58 months post-birth, respectively. For the annual estimation in Panels B, pre-birth is two calendar years before birth, we omit the year before and short-, medium- and longer-term refer to 0-1, 2-3 and 4 years after birth. Standard errors clustered at the mother level in parentheses. Significance levels: * < 10% ** < 5% *** < 1%.

Source: IEB, own calculations.

CHAPTER 4

Understanding Day Care Enrolment Gaps¹

For copyright reasons, this chapter is not included in the online version of this dissertation. It is published as *Jessen, J., S. Schmitz, and S. Waights (2020): Understanding day care enrolment gaps, Journal of Public Economics, 190, 104252*.

Available online at <https://doi.org/10.1016/j.jpubeco.2020.104252>

¹This chapter is joint work with Sophia Schmitz (Federal Ministry of Finance) and Sevrin Waights (DIW Berlin and CEP at the LSE). We are grateful to Mette Gørtz, Sandra McNally, Frauke Peter and C. Katharina Spiess, as well as seminar participants at the DIW Berlin for helpful comments. We also indebted to Benjamin Gedon for assisting us with the data. We thank the handling editor at the Journal of Public Economics, Christopher Walters, and two anonymous referees for excellent comments that helped to substantially improve the paper. Sevrin Waights acknowledges funding from a Marie Skłodowska-Curie Actions fellowship under the Horizon 2020 programme of the European Union (Project acronym: PAGE, Grant number: 752053).

CHAPTER 5

Center-Based Care and Parenting Activities¹

5.1 Introduction

Children in high-income countries are spending more time in child care centers than ever before: not only is the age of first entry decreasing, but the hours spent in daily care are also increasing over time (OECD, 2017). Nevertheless, parents are spending more time on activities with their children than they did in the 1960s (Dotti Sani and Treas, 2016).² Based on these trends, it is far from obvious that parenting activities are reduced as a result of the increasing usage of center-based care. Parental interactions play a key role in children’s development, independent of the role of learning institutions, such as child care centers (Del Bono et al., 2016; Cunha et al., 2006; Kim et al., 2018; Moroni et al., 2019; Todd and Wolpin, 2007). Specific activities, such as reading to the child, are particularly valuable (e.g. Kalb and van Ours, 2014; Mayer et al., 2019; Price and Kalil, 2019). Therefore, knowing whether center-based care increases, decreases, or has no effect on parenting activities is important for understanding the effects on child development (we discuss each possibility in section 5.3).

Our study asks: how do parenting activities differ when center-based care is being used? We estimate conditional differences with a rich set of control variables and provide bounds for our estimates. Although there is a huge economic literature on the

¹This chapter is joint work with C. Katharina Spiess (DIW Berlin and Freie Universität Berlin) and Sevrin Waights (DIW Berlin and CEP at the LSE). We are grateful to Jan Berkes, Nabanita Datta Gupta, Christina Felfe, Mette Gørtz, Ariel Kalil, Kevin Lang, Cheti Nicoletti, Björn Öckert, Alexandros Theloudis, Emma Tominey, Ulf Zölitz, and seminar participants at DIW Berlin, the Tenth International Workshop on Applied Economics of Education and at the Workshop of Labour and Family Economics (WOLFE) for helpful comments. Andrew Judy and Elena Ziege provided excellent research assistance. Sevrin Waights acknowledges funding from a Marie Skłodowska-Curie Actions fellowship under the Horizon 2020 programme of the European Union (Project acronym: PAGE, Grant number: 752053).

²Moreover, the type of parents who see the largest increases in parenting activities—i.e. more educated parents, according to Dotti Sani and Treas (2016)—are those who have seen the largest increases in usage of center-based care (e.g. see Jessen et al., 2018, for Germany).

effects of center-based care programs on various outcomes, evidence on the effect of center-based care on parenting activities is limited. Research mainly focuses on the effects of center-based care on maternal labor supply (for a recent overview see Müller and Wrohlich, 2020) and child development (e.g. Blanden et al., 2016; Cornelissen et al., 2018; Datta Gupta and Simonsen, 2012; Felfe and Lalive, 2018; Havnes and Mogstad, 2011b, 2015; Kuehne and Oberfichtner, 2020) with a few further studies looking at other outcomes like maternal well-being (i.e. Schmitz, 2020), child abuse (Sandner and Thomsen, 2020) and fertility (Bauernschuster et al., 2016). Studies by Baker et al. (2008) and Herbst and Tekin (2014) are some of the few economic studies to look at the effects of child care programs (in Canada and the U.S., respectively) on the style and quality of parental interaction (among other outcomes). However, while important, quality and style of parenting are not necessarily closely related to the time spent on parenting activities.

Depending on the context, many studies find positive effects of center-based care on child development, especially for children from less-educated parents, but some also show zero or even negative effects (e.g. Baker et al., 2019; Fort et al., 2020). The direction and size of the effect is most commonly thought to be related to the educational opportunities offered at the child care center relative to the home environment, with some studies focusing specifically on the role of center quality (e.g. Bauchmüller et al., 2014). However, this institutional channel typically takes the educational environment at home as a fixed consideration (e.g. Guryan et al., 2008; Kalil et al., 2012).³ A much less-explored channel is whether usage of center-based care might impact child development by changing the home environment, for instance, by affecting parenting activities.

Our main contribution is to use time-diary data to estimate effects of center-based care usage on parenting activities in Germany, a country with a universal child care system (see section 5.2). We do this by estimating conditional differences separately on (i) parents' overall time spent together with the child and (ii) the absolute amount of time spent on parenting activities.⁴ We also estimate effects for specific types of parenting activities such as reading and primary care. In doing so, we follow the child development literature, which distinguishes between activities that involve different levels of interaction (Fort et al., 2020; Kalil et al., 2012). We contribute to a very

³One of the few economic studies not to take the home environment as given is Kuger et al. (2019), which shows that the quality of center-based care affects the quality of the home environment, using established quality measures for both environments.

⁴Our analysis is restricted to families with one child below the age of ten for data reasons. Using a household survey, we show in a complementary analysis in Appendix section 5.A.5 that when we contrast results with and without the same sample restriction they hardly differ.

sparse literature addressing our question.⁵ To the best of our knowledge, the only existing economic study is Kröll and Borck (2013), which uses data from the German Socio-Economic Panel (SOEP) and finds that center-based care increases maternal interactions with children. However, the analysis is based on how often mothers report having undertaken specific activities with their children in the past fortnight, rather than precise time diary data. The few studies from other social sciences that examine the relationship between center-based care and parent-child interactions tend to find small decreases that come mostly through primary care rather than development-enhancing activities (e.g. Booth et al., 2002; Folbre and Bittman, 2004; Craig and Powell, 2013; Habibov and Coyle, 2014). However, these studies do not attempt to address selection on unobservables, and place little emphasis on the specific types of activities carried out.

Another major contribution of our study is to outline a framework of mechanisms and apply it to the data. We distinguish between direct effects, which are changes in parenting activities that occur while the child is at the child care center, and indirect effects, which are changes to parenting activities outside of center hours while the child is at home (e.g. in the evenings and on weekends). While direct effects are to a large degree mechanical, indirect effects indicate how parents respond in their parenting activities in the remaining time. Indirect effects may be either positive or negative depending on whether center-based care is a complement or a substitute for parenting activities, which itself depends on changes to parental motivation and time constraints. We apply the framework empirically by using the diary data to estimate effects on parenting activities at specific times of the day: during typical care center hours or outside of those hours. We explore whether the effects likely reflect changes to motivation or to time constraints by additionally looking at effects on non-parenting activities (such as paid work, housework and leisure). Previous studies neither distinguish between direct and indirect effects, nor attempt to systematically explore adjustment mechanisms.⁶ In doing so, our study contributes to a literature on the economics of parenting that tries to explain parenting decisions as rational choices that may be affected by the institutional environment (e.g. Doepke and Zilibotti, 2017; Doepke et al., 2019).

A further contribution is that we do not just focus on center-based care usage, *per se*, but on the effect of the dosage as well: We complement our main analysis with an

⁵Some studies focusing on the impact of maternal employment on parenting activities show that parental quality time with children does not need to decline with increases in maternal employment (e.g. Bastian and Lochner, forthcoming; Del Bono et al., 2016).

⁶A previous study that also analyzes the effect of center-based care on parenting activities also finds evidence of the impact of center-based care on non-parenting outcomes like housework (Craig and Powell, 2013).

examination of the effects of full-day vs. half-day care. We do this using the same time-use data and additional survey data, the German Family Panel (pairfam). The dosage of center-based care is an important margin since the literature finds quite differing effects on child development by hours of center-based care (e.g. Datta Gupta and Simonsen, 2010; Felfe and Zierow, 2018; Loeb et al., 2007).

Our method involves regressing time spent on parenting (and non-parenting) activities on an indicator for center-based care usage. We estimate an unconditional model, and a conditional model with a rich set of controls for child, parent, and household characteristics. To account for potential selection on unobservables into center-based care, we implement the coefficient stability approach of Oster (2019). Selection on unobservables is accounted for by assuming it relates to the degree of selection on observables, which itself is measured based on coefficient movements (and changes in the R^2) that occur when including control variables. We present ‘identified sets’ that estimate bounds based on assumed upper and lower limits for the degree of selection on unobservables. In general, we find that our coefficients are relatively stable to the inclusion of controls, thus suggesting fairly limited selection bias. We do not claim to estimate ‘causal’ effects, but are comfortable using the word ‘effect’ to describe our estimates since we believe them to be a decent proxy.

Our estimates imply that center-based care usage reduces the overall time that parents spend with their child but that there are only small effects on the time spent on parenting activities with the child and on educational activities, specifically. Our results are consistent with the few existing studies that find only small decreases in parenting activities (see above). Indirect effects center closely around zero. These effects combined with an increase in paid work and a reduction of sleep during non-center hours, suggest an increase in parental motivation.

Finally, our results show that full-day care, in comparison to half-day care, is also associated with a decrease in parenting activities. We additionally find decreases in the frequencies of certain parenting activities, although the effect sizes are small. This is in line with the literature that finds more limited child development effects at this margin (e.g. Felfe and Zierow, 2018).

5.2 Institutional background

In 2020, 35 percent of children in Germany under three and 93 percent of those aged three to five were enrolled in center-based care. For both age groups, just over half of the enrolled children were in full-time care, defined as 35 hours or more per week. The child care system in Germany can be characterized as a virtually universal, strongly state-subsidized system. For-profit providers play a very limited role, with only 2.6 percent of

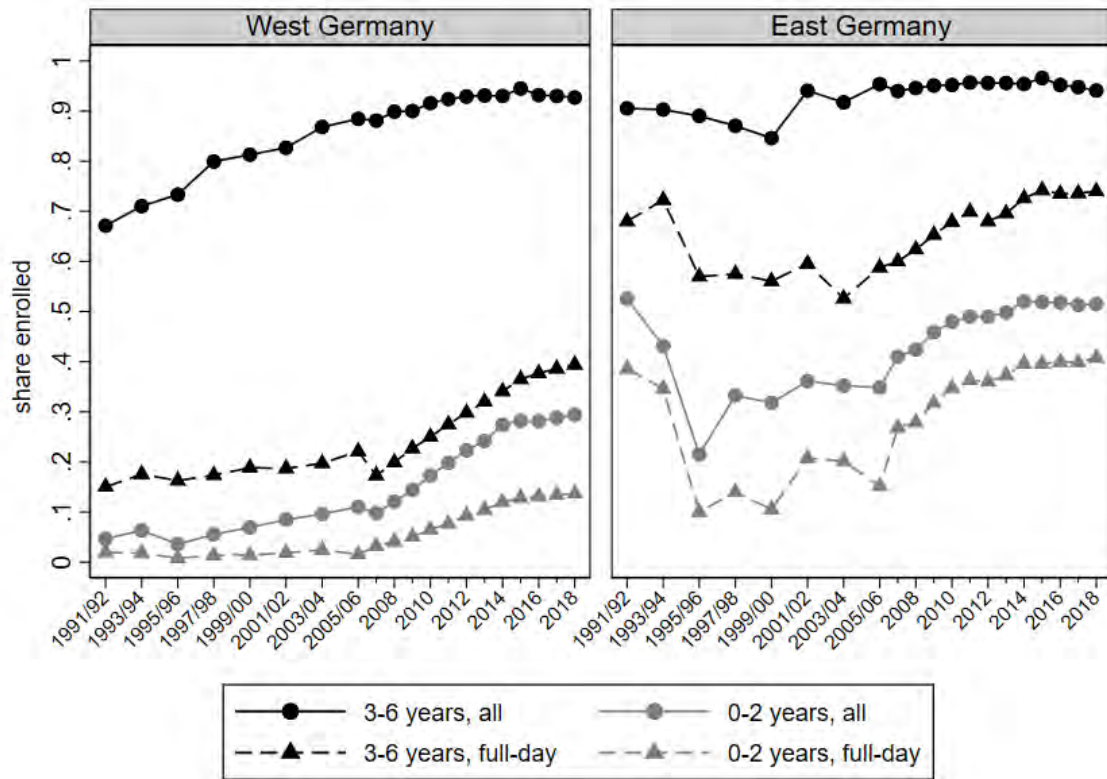
institutions in 2017 being private and non-charitable (Destatis, 2017). Parental fees are mostly income-dependent and relatively low compared to most other OECD countries (OECD, 2020), with many states having even abolished fees altogether for older age groups at least (Huebener et al., 2020). In 2012, average fees amounted to 144 Euros per month and family, on average (Schroeder et al., 2015). In general, parents cannot obtain higher quality by paying higher fees, which weakens the link between family income and center-based care quality compared with countries using a market-based system (Stahl et al., 2018).

Figure 5.1 shows enrollment rates in center-based care for under and over three year olds separately for East and West Germany over the time period covered by our analysis. For over-threes, the majority of the expansion in child care center slots took place in the 1990s in response to the 1996 introduction of a legal entitlement to a place for children over three years and a general trend in Europe to expand center-based care for children three years and older (see e.g. Spieß, 2008). In both East and West Germany, enrollment rates for over-threes have been above 80 percent since before 2000. Despite a strong increase in full-day enrollment in West Germany in the 2010s (Jessen et al., 2018), full-day rates remain below 50 percent. In East Germany, full-day enrollment rates are much higher, covering 74 percent of over-threes children in 2018.

In contrast, for under three year olds, enrollment rates were very low well into the mid-2000s, particularly in West Germany. In 2008, a federal law (KiföG) was passed, extending the legal claim to a place at a child care center to children of at least one year of age, coming into effect in 2013. The legal change and the accompanied increased provision came in response to a long-lasting over-demand for center-based care, in particular by parents with infants and toddlers (i.e. Spieß and Wrohlich, 2005; Wrohlich, 2008). However, while enrollment rates for under-threes subsequently climbed, demand increased further still resulting in a continuation of shortages (Jessen et al., 2018).

Parents in Germany make frequent use of informal care, especially by grandparents. In 2017, between 50 percent and 60 percent of all children from six months old until the age of six years had grandparents as caregivers; for older children, grandparents were mainly used in addition to center based care. Other private caregivers looked after between only 10 percent and 30 percent, of children, depending on child age. Nevertheless, informal care, such as that offered by grandparents, is typically for only a few hours per week and complementary to formal care. This is shown in Appendix Tables 5.A.4 and 5.A.5: informal care is higher for children who attend a child care center, suggesting that informal care may have been used to extend hours of formal care, rather than to substitute it.

Figure 5.1: Share of children enrolled in center-based care by region, age group, and time



Note: Figure shows the share of children aged 0-6 years enrolled in center-based care and in full-day care by region (West vs. East Germany) and age group over time. enrollment includes formal child care centers and care by qualified publicly funded child minders. Data for 1991/92-2005/06 from the German Socio-Economic Panel (SOEP v35), which is a long-running household survey containing information on about 15,000 households per year (Goebel et al., 2019). For precision, data is pooled in two-year bins. Annual statistics since 2007 from the German Federal Statistical Office (starting that year, official administrative data contain the share in full-day care).

Parental care in Germany is characterized by a strong gender divide, with mothers acting as the primary caregivers (Schober, 2014). Parenting activities (and housework), therefore, are carried out to a much larger degree by mothers despite a slight narrowing of the gender gap since the 1990s, as illustrated in Appendix Figure 5.A.3. Consistent with the ‘primary-male-breadwinner’ model, evidence shows that the roll-out of center-based care, as described above, had an employment effect for mothers but made no difference for fathers (Müller and Wrohlich, 2020). In addition, parenting activities (in minutes per day) exhibit a strong upwards time trend for both mothers and fathers, which is broadly comparable to that found in other countries (Dotti Sani and Treas, 2016).

5.3 Adjustment mechanisms

This section discusses ways in which center-based care usage might affect parenting activities. We focus on the amount of time spent on activities as an outcome rather than any measure of parenting quality or style. We define *direct effects* as changes that occur during the time that the child spends at the child care center, and *indirect effects* as changes that occur outside of center hours as a result of parental adjustments.

The direct effect (i.e. during center hours) of center-based care on parenting activities will be negative if center-based care usage reduces the time that a parent spends with their child, when they would have otherwise engaged in some parenting activities in the counterfactual. No effect is only possible if center-based care fully crowds out informal care arrangements, by grandparents, for instance, or if despite being with the child, no parenting activities are done in the counterfactual. As the direct effect is largely mechanical, we will put less emphasis on this effect in the discussion of the results.

Indirect effects (i.e. outside of center hours) may occur as follows:

- Positive: if center-based care is a complement to parenting activities. This could be if center-based care reduces parental time constraints or increases parental motivation to interact with their child. Time constraints may be reduced if parents use the center-based care hours to complete other tasks, such as paid work or housework, thereby freeing up non-center hours for parenting activities. Furthermore, not being at home with a child may mean there is less cleaning and tidying to be done in the evening.⁷ Motivation may be increased if spending less time with the child overall means that parents try to ensure that they do more activities with the child in the remaining time. Further, it could be that center-based care inherently encourages parents to interact with their child, e.g. through teacher recommendations (see e.g. Cornelissen et al., 2018; Kuger et al., 2019).⁸ Moreover, if center-based care has a direct effect on children's cognitive or socio-emotional development, parents could adjust their inputs in response to this and increase their time spent on specific parenting activities (see Nicoletti and Tonei, 2020).

⁷One thing to note is that if increased activities are due to a reduction of time constraints, then this may reflect lower parental stress and a higher quality of interaction than captured by a simple increase in parenting activities. Sandner and Thomsen (2020) find evidence that the expansion of center-based care in Germany led to a reduction in cases of child abuse and neglect. They propose a reduction of mental and physical overburdening of parents as the driving mechanism underlying this. Additionally, Schmitz (2020) finds that provision of public child care in Germany directly increases maternal well-being.

⁸This holds especially true if care center staff observe developmental deficiencies, if they believe that educational activities are performed too rarely and/or if they believe that parents are unaware of the benefits associated with them.

- Negative: if center-based care is a substitute for parenting activities. This could be the result of a decrease in parental motivation, e.g. if parents feel that certain activities are no longer necessary since they are already done with their child in center-based care. This might be the case in particular if there is a notable positive effect of center-based care on child development. Furthermore, substitution could occur through a worsening of parental time constraints, e.g. if parents use center-based care hours to take on significant extra activities, such as paid work, meaning they have more tasks to do in the evenings instead of parenting activities.
- No effect might arise if center-based care is neither a substitute nor a complement, i.e. if there are no motivation and time constraints effects or if they are counterbalanced.

While we have priors for the direct effects, there is little evidence on which to base hypotheses regarding the direction of indirect effects. A separate question is what direction the overall effect might be (i.e. direct and indirect together). There might be positive indirect effects on parenting activities that are large enough to overcompensate for a negative direct effect. Again, we have little guidance to form any priors in this regard. In Appendix section 5.A.1, we provide some stylized examples to further illustrate the mechanisms with specific cases.

5.4 Data and empirical approach

5.4.1 German Time-Use Survey

We use diary data from three waves of the German Time-Use Survey, which is a repeated cross-section of around 5,000 households per survey wave taken in 1991/92, 2001/02 and 2012/13 (Destatis, 2015; Maier, 2014). The diary data record the main and (optional) secondary activity of each adult household member in five- or ten-minute slots over two or three days using a three-digit classification (see Appendix Tables 5.A.2 and 5.A.3 for further details).⁹ An example of a three-digit activity is ‘reading to child’, which is from the two-digit activity of ‘child care’, which belongs to the one-digit category of ‘work in the household’. We use the activities recorded under ‘child care’ as our parenting activities. In addition to recording specific activities, the survey also indicates for each time slot whether it was spent with a child under the age of ten years present. Importantly, the parent need not necessarily record a parenting

⁹The first wave consists of two successively recorded days that are uniformly distributed, meaning that about three quarters of the days in the sample are weekdays. In the two later waves, individuals’ activities are recorded over three days, two weekdays and one weekend day.

activity as the main or secondary activity while spending time with the child.¹⁰ The data also includes information on households—such as usage of center-based care, age of youngest child, number of children, single-parent household, and location in East or West Germany. At the respondent-level, the data includes information on age, gender, education, marital status, and economic activity.

We use parent-days as the unit of observation for our analysis. We define two main measures of parental involvement: (1) time with child, as the number of minutes on a day that a parent spends together with their child; and (2) parenting activities, as the number of minutes on a day spent on child care activities as the main activity. We think of time with child as capturing a more basic form of child care than parenting activities, since the latter involves specific interactions with the child, which may better foster child development (see, e.g., Kalil et al., 2012). Thus, we think of parenting activities as being the relevant measure of the educational potential of the home environment. In some analyses, we also distinguish between particular types of parenting activities: reading to the child, playing with the child, talking with the child, and primary care.¹¹ We also estimate effects on non-parenting activities, like ‘paid work’, ‘housework’, and ‘leisure’ to investigate mechanisms.¹²

In our analysis, we look at effects for mothers and fathers separately and we differentiate households by maternal education. We follow the common practice in the literature (e.g. Fiorini and Keane, 2014; Nicoletti and Tonei, 2020) of grouping households by maternal education, both because it is highly correlated with paternal education and also because mothers are usually the primary caregivers in our context.¹³ We define the educational background as higher if the mother (or, very rarely, male single parent) in the household holds a secondary school certificate from the upper educational track in Germany, which ends with a university entry degree (Abitur). The education split is motivated by differential effects of center-based care on child development found in the literature and well-established differences in parenting activities by education (see e.g. Bradley et al., 2001; Dotti Sani and Treas, 2016; Gimenez-Nadal and Molina, 2013; Guryan et al., 2008; Kalil et al., 2012). We also differentiate by the time of the day (center hours or non-center hours) in specifications that aim to estimate direct and

¹⁰For example, a parent may record ironing as the main activity and watching television as the secondary activity, while also indicating that the time was spent with a child.

¹¹‘Primary care’ covers bodily hygiene, feeding and clothing the child, as well as passive supervision (i.e. ‘keeping an eye on’ the child).

¹²Leisure consist of the 1-digit activities ‘social life and entertainment’, ‘sport, hobbies and games’, and ‘media usage’ as shown in Appendix Table 5.A.2 for the 2012/13 survey wave.

¹³In fact, women spent more hours per day on child care than men in all European countries analyzed in Gimenez-Nadal and Molina (2020).

indirect effects. In these specifications the outcomes are the sum of minutes dedicated to each activity during either center hours or non-center hours in a day.

We restrict our sample to parents whose youngest child is of the enrollment age for center-based care (i.e. under six years old). Furthermore, we drop all parents who have more than one child under ten years old. This restriction reduces the sample by 58% but ensures that time with child measures effects on the enrolled child and not any potential indirect effects on time with an older child (who is also under ten years). We do not expect the effects to be dramatically different for the dropped households (with further children under ten) since it is enrollment of the youngest child in center-based care that usually makes the key difference in terms of the child care responsibilities of parents. Indeed, in Appendix section 5.A.5 we use the household survey data that reports activities on a child basis to show that the coefficients for children of the relevant age (three to six years) are similar whether or not we make the one child under ten restriction. After these restrictions, the main sample comprises 4,490 parent-days and 1,818 person observations. Table 5.1 presents summary statistics of the main sample, split by enrollment in center-based care. Differences between the samples are apparent; parents of children enrolled are on average older, have obtained more education, are more likely to be economically active and more reside in East Germany. Children in center-based care also are older. In the empirical analysis it is thus paramount to control for exogenous unbalanced characteristics, as these are also likely to correlate with parenting behavior.

To illustrate the diary data, Figure 5.2 plots the number of minutes per hour of the day spent doing different activities by usage of center-based care. In these descriptive plots we focus only on mothers observed on weekdays, since this subsample of parent-days demonstrates the clearest differences in terms of direct and indirect effects.¹⁴ However, in our analysis, the baseline specifications pool mothers and fathers as well as weekdays and weekend days, to give a clearer picture of effects on parenting activities overall. The descriptive plots show that center-based care is associated with mothers spending less time with their children on weekdays during regular care center hours (08:00-16:00), especially in the morning. There is also a lower share of parenting activities, although it is less pronounced and followed by an apparent increase in the late afternoon and evening (16:00-20:00). Time in paid work is higher for center-based

¹⁴Appendix Figure 5.A.2 shows a version pooling mothers and fathers on all days. Appendix Figure 5.A.3 shows the average daily duration of the activities shown in Appendix Figure 5.2 separately for fathers and mothers by sample wave. Note that the decrease in time spent in paid work by mothers after the first sample wave (1991/92) is driven by mothers in East Germany in the aftermath of the German reunification.

Table 5.1: Sample characteristics by enrollment in center-based care

Variable	(1)	(2)	(3)
	Center-based care		Difference
	No	Yes	
<i>Parent characteristics</i>			
Female (0/1)	0.54 (0.02)	0.55 (0.01)	0.006 (0.024)
Age in years	33.37 (0.26)	36.27 (0.21)	2.904*** (0.332)
Higher educated (0/1)	0.37 (0.02)	0.41 (0.01)	0.048** (0.024)
Married (0/1)	0.82 (0.02)	0.79 (0.01)	-0.030 (0.019)
Single parent (0/1)	0.08 (0.01)	0.09 (0.01)	0.016 (0.014)
Economically active (0/1)	0.57 (0.02)	0.66 (0.01)	0.089*** (0.024)
Economically part-active (0/1)	0.10 (0.01)	0.15 (0.01)	0.053*** (0.016)
East Germany (0/1)	0.13 (0.01)	0.36 (0.01)	0.237*** (0.019)
<i>Child characteristics</i>			
Girl (0/1)	0.49 (0.02)	0.52 (0.01)	0.024 (0.024)
Age in years	1.30 (0.05)	3.61 (0.04)	2.310*** (0.060)
<i>Sample period (row shares)</i>			
1991/92	0.36	0.63	
2001/02	0.39	0.60	
2012/13	0.29	0.71	
Person-day observations	1588	2902	4490
Person observations	647	1171	1818

Robust standard errors in parentheses.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

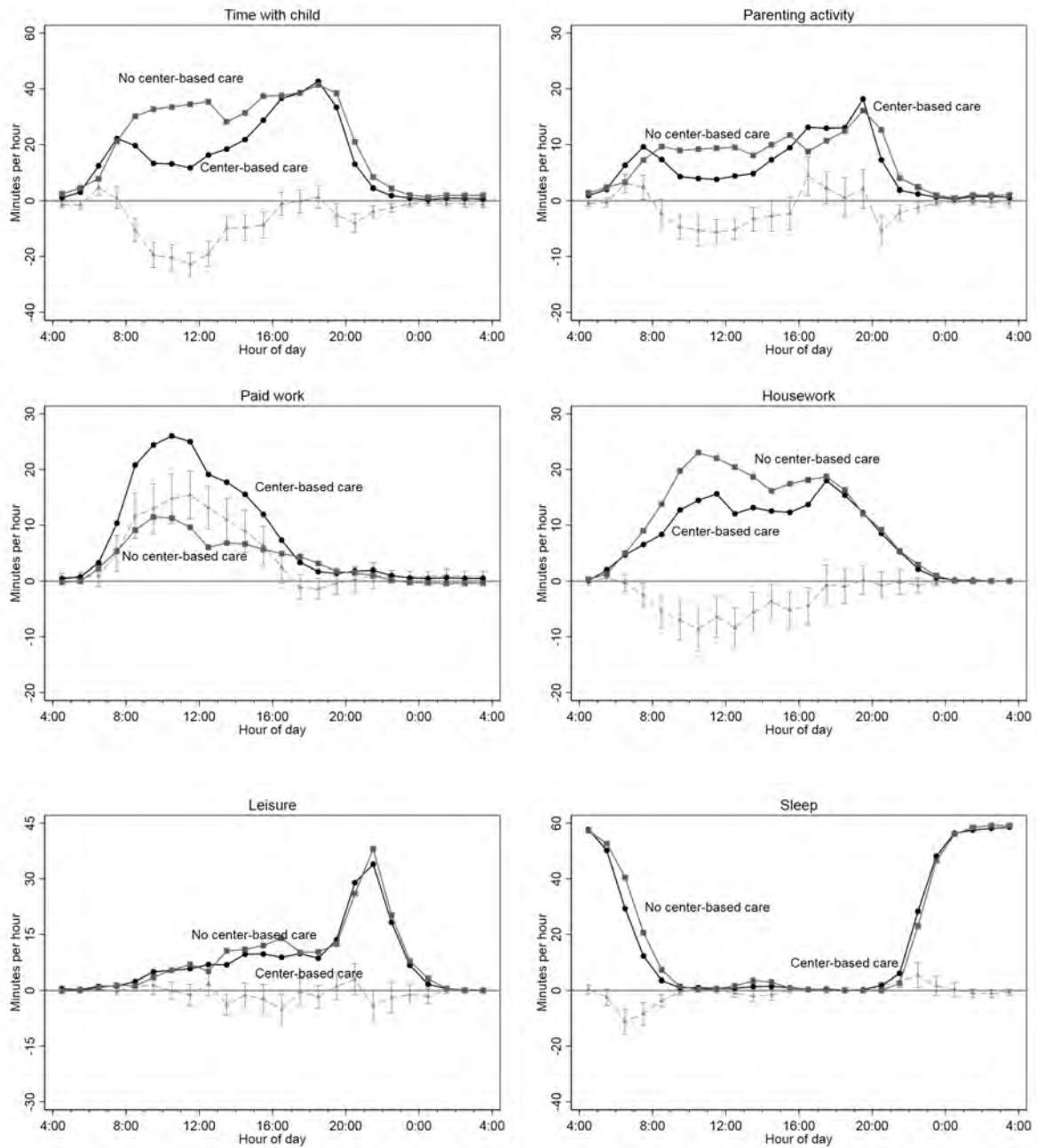
care users, while both housework and leisure are lower during center-based care hours. Finally, we see that mothers with their child in center-based care wake up earlier.

5.4.2 The German Family Panel

For an additional analysis of the effects of full-day vs. half-day center-based care, we use the German Family Panel (pairfam), which is a longitudinal household survey collected annually since 2008 and used for researching partnership and family dynamics.¹⁵ The survey records the frequency of specific parenting activities, but only for older children

¹⁵See Brüderl et al. (2020) for a data documentation.

Figure 5.2: Mothers' activities on weekdays by usage of center-based care



Notes: Circles denote mothers with a child in center-based care, squares those without. Differences and averages are estimated in weighted regressions with indicators for child age and evaluated at mean values. Whiskers indicate 95% confidence intervals. Data consists of time slots in ten minute intervals (five in the first survey wave), which then are aggregated by hour of day. Sample includes mothers on weekdays only. Appendix Figure 5.A.2 shows the same plots but with fathers and mothers pooled over all days. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

(three years and above) and only since 2013. Therefore, we restrict our sample to data between 2013 and 2019 for children between three and six years of age. While

the survey does not collect precise diary data, it gives us around ten times as many observations as does the time-use sample for the full-day vs. half-day care analysis, allowing for greater precision in estimation. We cannot use pairfam to examine day-care vs. no day-care since the activity-questions are only available for children aged three and above who nearly all attend center-based care.

For each child of a parent, the survey asks: *How often have you done the following things with your child during the past 3 months?* An overview of frequencies of shared activities for children in half-day and full-day center-based care is shown in Appendix Figure 5.A.4. We code indicator variables for whether each activity is carried out at least daily as outcome variables. The data also include information on the type of care each child uses, as well as parent, child, and household characteristics. We code children as being in full-day care if they are in center-based care in the morning and afternoon and half-day if they are at center-based care in the morning or afternoon. Appendix Table 5.A.1 shows summary statistics for children attending half-day or full-day care and for their families and households.

5.4.3 Empirical approach

We start our analysis with an unconditional model, whereby we regress parenting activities on center-based care usage controlling only for a set of indicators of child age in years.¹⁶ This unconditional model corresponds to the daily sum of the differences plotted in Figure 5.2. Next we estimate a conditional model that accounts for selection into center-based care based on observable characteristics: child age indicators, child gender, parent age, parent age squared, parent gender, parental education indicators for secondary school track (upper, middle, or lower) and for university degree, marital status, single parent status, number of children in household, as well as an indicator for weekday observations. We also include indicators for survey wave \times region (East or West Germany) to control for the different institutional settings described in section 5.2. Despite having a fairly rich set of controls, it remains possible that selection into usage of center-based care is driven to some degree by unobservable parent characteristics that are also correlated with their parenting activities. Due to this, the estimates ought to be interpreted as conditional correlations rather than as 'clean' causal effects.

In an additional step, we account for possible selection on unobservables by examining coefficient stability across unconditional and conditional models. We follow Oster (2019) in making assumptions regarding (i) the maximum achievable R^2 , i.e. R_{max}^2 , and

¹⁶Given near zero usage rates in the first year of center-based care in this age range and near full usage in the last two years, the relationship between child age and usage has a relatively large deterministic component. Therefore, we include it in the unconditional model. Similar choices are made by Oster (2019) to include, for example, weeks of gestation in the unconditional model for birth weight.

(ii) the extent of selection on unobservables relative to selection on our set of included controls, i.e. δ . Our main specification assumes $R_{max}^2 = 1.3\tilde{R}^2$ where \tilde{R}^2 is the R^2 of the conditional model.¹⁷ We assume that δ is bounded such that $\delta \in [0, 1]$. At the most ‘optimistic’ bound of $\delta = 0$ there is no selection on unobservables. At the most ‘pessimistic’ bound of $\delta = 1$, selection on unobservables plays an equal role to selection on the included controls. This seems a reasonable upper bound given we have a fairly rich set of controls. The corresponding identified set of estimates gives us the upper and lower bound for the true effect assuming that the real δ falls between the two extremes. Whereas the bounds presented assume that selection on unobservables follows the same direction as selection on unobservables, we also provide the δ that would be required based on the coefficient movements and R_{max}^2 for the true coefficient to be zero. A large *absolute* value of δ here indicates that the true coefficient is zero only if selection on unobservables is very large relative to selection on our controls. Relative selection on unobservables may go in the same or opposite direction as observables.

5.5 Results

5.5.1 Effects of center-based care usage

First, we describe the effect of center-based care usage on parenting activities. Table 5.2 reports the results for all parents together, for mothers only, and for fathers only, each differentiated by the educational attainment of the mother in the household. For the group of households with lower maternal education, center-based care reduces time with child for both mothers and fathers (by 113 minutes and 30 minutes, respectively, in the conditional model). The reductions in parenting activities by comparison are relatively small (around 13 minutes for mothers and no significant difference for fathers). For households with higher maternal education, conversely, center-based care reduces time with child for mothers only (and by a smaller amount, 70 minutes, compared with mothers from lower education households). The reductions in parenting activities are a little larger in absolute size compared with households with lower maternal education, and especially so as a share of the reduction of time spent with the child, suggesting a more parenting-rich environment in the no center-based care counterfactual.

Comparing the conditional and unconditional models shows that the coefficients do not change by much upon adding control variables. Nevertheless, the controls have decent explanatory power as indicated by the substantial increase in the R^2 between

¹⁷The value of 1.3 has been derived by Oster (2019) through examining under which assumptions of R_{max}^2 90% of randomized results examined in her study survive.

Table 5.2: Effects of center-based care on parenting activities

	Households with lower maternal education			Households with higher maternal education		
	All parents (1)	Mothers only (2)	Fathers only (3)	All parents (4)	Mothers only (5)	Fathers only (6)
<i>Outcome: Time with child (minutes per day)</i>						
Unconditional	-86.3*** (12.9)	-141*** (16.1)	-29.2* (16)	-44.4*** (15.1)	-81.6*** (18.2)	-.157 (20.8)
Conditional	-74.4*** (11.5)	-113*** (16.6)	-30.3** (14.9)	-39.9*** (13.3)	-70.3*** (18.1)	-4.8 (18.9)
Mean	336.475	427.119	227.138	360.388	444.836	258.904
Identified set	[-74.390, -69.609]†	[-113.481, -91.122]†	[-30.682, -30.281]†	[-39.895, -37.911]†	[-70.345, -60.650]†	[-6.729, -4.797]†
δ for 0 coefficient	8.678	3.083	33.444	9.832	3.964	-3.004
R^2 (unc., con.)	(0.037, 0.353)	(0.037, 0.238)	(0.037, 0.286)	(0.037, 0.326)	(0.037, 0.219)	(0.037, 0.280)
Observations	2482	1357	1125	2008	1096	912
<i>Outcome: Parenting activities (minutes per day)</i>						
Unconditional	-8.96** (4.49)	-17.7*** (6.27)	.311 (4.93)	-12.5** (6.28)	-18.3** (8.4)	-5.91 (7.9)
Conditional	-6.29 (4.26)	-12.9** (6.45)	2.62 (4.98)	-12.4** (6.19)	-15.6* (8.95)	-9.9 (8.27)
Mean	90.520	125.380	48.471	106.688	140.370	66.212
Identified set	[-6.289, -4.637]†	[-12.878, 4.597]	[2.620, 3.774]†	[-12.397, -12.313]†	[-15.627, 1.088]	[-13.466, -9.895]†
δ for 0 coefficient	3.201	0.797	-2.754	11.372	0.964	-4.235
R^2 (unc., con.)	(0.098, 0.347)	(0.098, 0.352)	(0.098, 0.149)	(0.098, 0.301)	(0.098, 0.278)	(0.098, 0.169)
Observations	2482	1357	1125	2008	1096	912

Notes: Table shows coefficients from OLS regressions of the outcome variables on an indicator variable for usage of center-based care. Unconditional coefficients are from a regression that includes only indicators for child age in years. The conditional coefficients are from regressions that include the child age dummies, and additionally child gender, parent age (linear / squared), parent gender, parental education indicators for upper, middle, or lower secondary school track (upper drops out for sample of mothers due to split by maternal education) and for university degree, marital status, single parent status, number of children in household, a weekday indicator, and wave \times region indicators. Households with higher maternal education are where the mother in the household (or single parent) was in the upper secondary school track (required to enroll in university) and those with lower educated mothers are where the mother took the lower or middle track. The identified set shows coefficients obtained using the method developed by Oster (2019), where $R_{max}^2 = \min \{1.3 \times \bar{R}^2, 1\}$ assuming selection on unobservables is between zero ($\delta = 0$) and a level equal to selection on observables ($\delta = 1$). † denotes that the identified set excludes zero. The δ for 0 coefficient row shows for each outcome variable shows how large the relative selection on unobservables must be to obtain a coefficient of 0. R^2 (unc., con.) shows the R^2 of the unconditional and conditional regressions. Robust standard errors reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

the models.¹⁸ Since the coefficients are fairly stable to the inclusion of these important controls, we end up with identified sets that suggest relatively tight ranges when accounting for potential selection on unobservables. While center-based care usage is related to certain observable characteristics (evident in Table 5.1), the stability of the coefficients in Table 5.2 suggests that these differences are not, on average, associated with very different patterns of time use. Only for parenting activities of mothers the identified sets narrowly include zero.

In Appendix section 5.A.4, we explore heterogeneities for mothers and fathers beyond the household split by maternal education. We find that the negative effects on both time with child and on parenting activities increase by survey waves,¹⁹ consistent with the increasing time that children spend in center-based care over the period. The effects are also greater during weekdays, as one would expect, but there does appear to be some spillover to the weekends, thus justifying the pooling of these observations for the

¹⁸The increases in R^2 are around 0.2 on average between the samples. This is in the upper end of the distribution of the studies examined in Oster (2019).

¹⁹In fact, we identify no negative effects on parenting activities in the first two survey waves.

main analysis. We also see that parenting activities are not reduced for girls, which is driven by an increase of mothers' interactions with daughters outside of care hours. This is consistent with research for the U.S., Canada and the UK showing evidence of more interaction in same-sex parent-child relationships (e.g. Baker and Milligan, 2016; Lundberg et al., 2007), except that our result relates to changes in activities from center-based care usage.

5.5.2 Direct and indirect effects

In order to explore the mechanisms, Figure 5.3, panel (a), plots estimates (identified sets and 90 percent confidence intervals) by time of the day and maternal education (circles for lower, squares for higher maternal education) for parenting outcomes. The effects during typical care center hours (8am-4pm on weekdays) aim to capture direct effects, whereas changes during non-center hours (all remaining hours, i.e. 4pm-8am on weekdays, and full weekend days) reflect indirect effects. In Appendix Table 5.A.7, we present the full regression table, which also includes separate effects for the 'night time' (which we define as 8pm-8am).

The figures illustrate that both direct and indirect effects play an important role in explaining the differences in effects between households with lower and higher maternal education. During center hours, mothers and fathers in households with lower maternal education reduce their time with child by more than mothers and fathers in households with higher maternal education (though the differences are not statistically significant). Also during non-center hours is the time with child reduced in lower education households, which is not found for higher educated ones. The reduction in parenting activities comes entirely from center hours, with no significant adjustment effects observed during non-center hours. I.e., despite lower education mothers spending less time with their child, they manage to keep parenting activities constant.

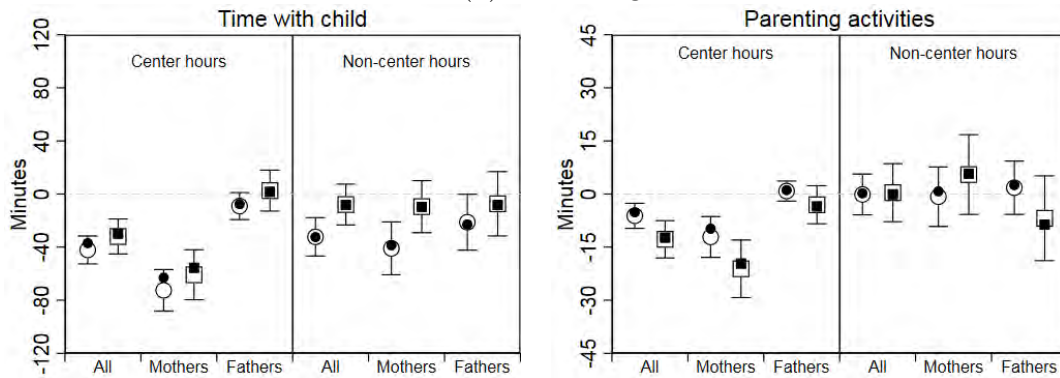
In Appendix Table 5.A.8, we additionally show which specific child care activities are affected:²⁰ for households with less educated mothers, there is an increase in reading to the child outside of center hours and a decrease in primary care at night-time. For the higher maternal education households, primary care on evenings and weekend days is increased.

Figure 5.3, panel (b), presents the effects of center-based care for four non-parenting activities: paid work, housework, leisure, and sleep by time of day and maternal education. The figures show that paid work increases during center hours (a direct effect) are entirely driven by mothers, with effects that are a little smaller in size to the decreases in time with the child. Further, there are decreases in housework that pre-

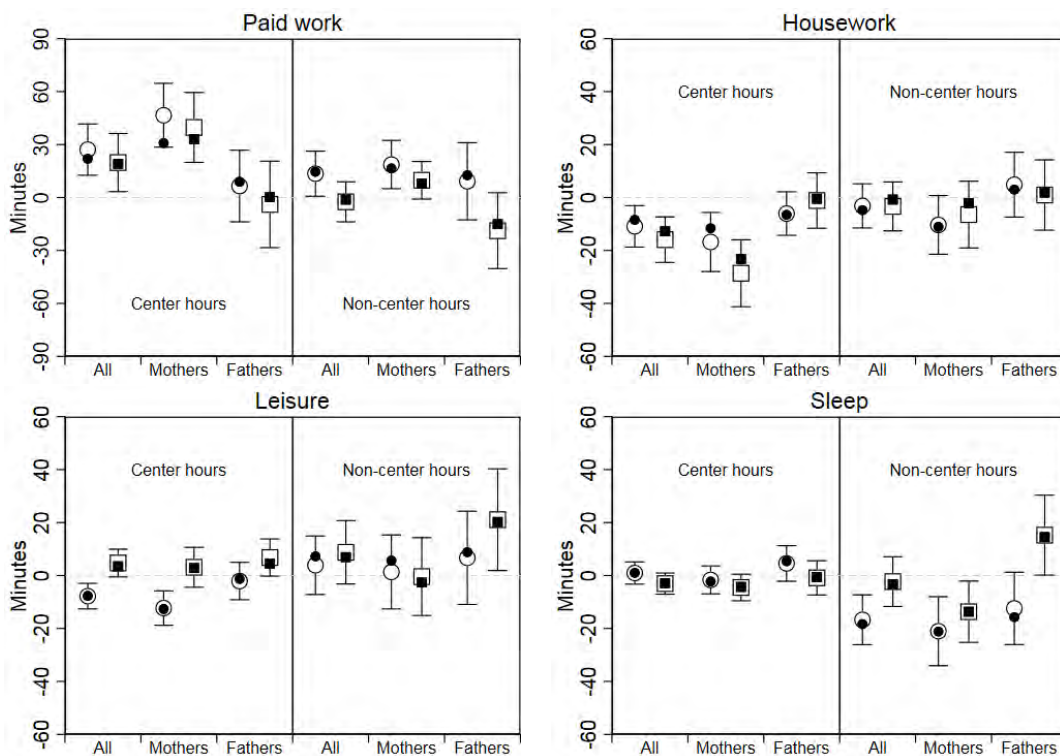
²⁰Those specific child care activities are not available in the 1991/92 wave of the survey.

Figure 5.3: Effects by time of day and education

(a) Parenting



(b) Other activities



Notes: Center hours are from 8am-4pm on weekdays, non-center hours are the remaining hours on weekdays (12am-8am and 4pm-12am) and the entire weekend days. Circles denote lower education households, and squares denote higher education. Education level of the household is based on whether the mother in the household (or the single parent) has a secondary school degree from the higher track (*Abitur*). The plots show the conditional difference in outcome variables by center-based care usage. Each estimate is based on a separate regression of the outcome summed over center hours or non-center hours on an indicator for usage of center-based care and controls (see notes to Table 5.2 for details) using all three waves of the time-use survey (1991/02, 2001/02 and 2012/13). The hollow shapes and whiskers indicate the conditional coefficient ($\delta = 0$) and the 90% confidence intervals. The filled shapes indicate estimates under the assumption of $\delta = 1$, i.e. equally large selection on unobservables as on observables. The filled and hollow shapes together indicate the identified set. Appendix Table 5.A.7 reports coefficients along with means of the outcome variables, and the δ required for zero coefficient, as well as separating out effects occurring at ‘night’ (which we define as 8pm-8am). Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

sumably would have been done during time with the child had it been at home. This is consistent with evidence that mothers use day-care to take up paid work (Müller and Wrohlich, 2020) instead of multi-tasking child care and housework. Mothers from

lower education households also experience a reduction in leisure during center hours, an indication of worsened time constraints, whereas higher educated mothers' leisure is not affected.

Turning to non-center hours (indirect effects), the figures provide an insight into whether the negligible effects on parenting activities are driven by effects on parental motivation or by changes to parental time constraints. For the lower maternal education households, there is an increase in paid work outside of center hours that matches a decrease in sleep. The differences are a little larger for mothers than for fathers and may reflect early shifts (before 8am) that require earlier waking.²¹ For lower education mothers there is a decrease in housework outside of center hours. This decrease in housework may explain how lower education mothers carry out just as many parenting activities despite spending less time with the child outside of center hours: the time before the child goes to bed is more concentrated on activities with the child rather than on other activities such as housework. This could reflect either a motivation effect (e.g. a change in priorities) or a time constraints effect (e.g. if there is less housework to do). For the higher maternal education households, there is also an increase in paid work and a decrease in sleep for mothers (albeit less pronounced) but for fathers the differences go in the opposite direction: a decrease in paid work and an increase in sleep outside of center hours. This result suggests a potential easing of time constraints for fathers from households with higher maternal education. Nevertheless, these fathers increase their leisure time outside of center hours suggesting that potential positive motivation effects may be playing less of a role than for other parents.

5.5.3 Effects of full-day vs. half-day center-based care

Thus far, the analysis has focused on the effects on parenting activities of using center-based care compared with not using it, irrespective of the number of hours of care used per day. The full-day vs. half-day margin may have different effects on parenting activities, which we explore in this section. Knowing the effects of full-day care on parenting activities is important since this has become the relevant decision margin for many parents (children over three years in Germany nearly all attend center-based care—see Figure 5.1). It may also contribute to our understanding of the child development effects for full-day care, which tend to be less beneficial for children from lower SES households (e.g. Loeb et al., 2007; Felfe and Zierow, 2018)

The 2012/13 wave of the German time-use survey contains information on the hours of center-based care normally used. Figure 5.4 plots the full-day vs. half-day effects (i.e.

²¹Appendix Table 5.A.7 reveals that these changes occur during the 'night' which is defined as 8pm to 8am.

conditional on usage of center-based care) on parenting (panel a) and non-parenting activities (panel b). As before, we plot estimates by time of day for both parents, mothers and fathers, and by maternal education. Both the higher and lower maternal education groups see increases in time with child during non-center hours, and small decreases in the parenting activities during center hours with no change outside of center hours.

The decreases in time with child during center hours for higher education mothers coincide with changes to paid work and housework, as before. However, in contrast to usage vs. non-usage, the increase in paid work are similar in size to the reductions in time with child suggesting worsening time constraints.

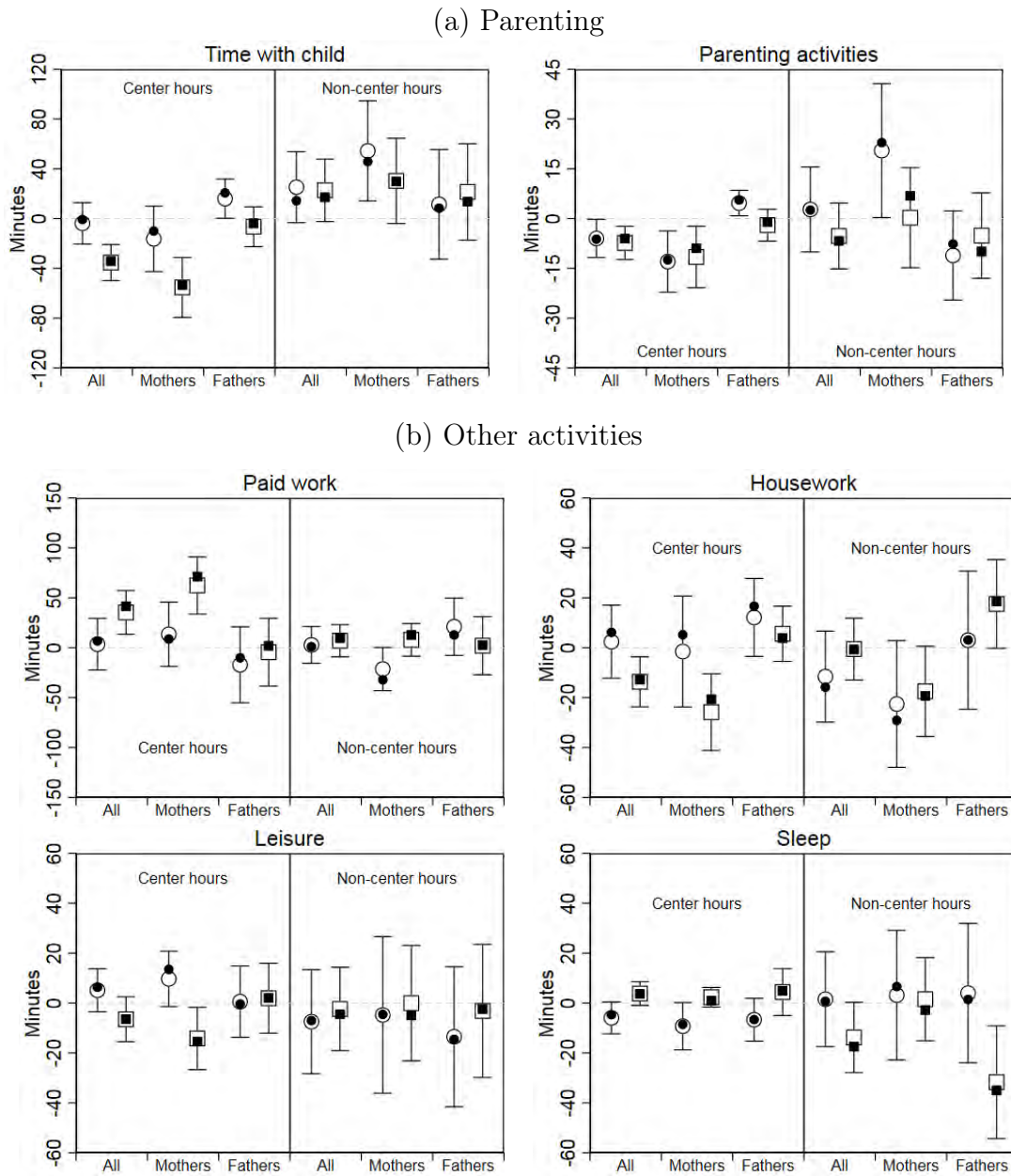
In order to investigate full-day effects with greater precision, along with effects on specific parenting activities, we turn to the German Family Panel (pairfam).²² Using this data, we estimate effects of full-day vs. half-day care on the probability of carrying out specific parenting activities on at least a daily basis. Table 5.3 shows the effects of full-day care on specific parenting activities (Panel A). We think of the first four activities (reading, music, art, and playing) as educational activities and the last three (outdoors, sports and TV) as recreational activities. For households with lower maternal education,²³ reading and playing is negatively affected by usage of full-day care, with effects being mainly driven by fathers. Music and arts are unaffected for mothers and fathers in lower educated households. In higher educated households, in contrast, reading is not reduced, but negative effects for arts and playing come from mothers. Finally, the frequency of musical activities is not reduced for either household type. For recreational activities, daily outdoor activities become less likely with full-day care, but sports and TV are unaffected. The negative effects for higher educated parents come through mothers rather than fathers.

Consistent with the time-use data, full-day care also allows for an increase in paid work that is larger for mothers with higher educational attainment (Panel B) and non-existent for fathers from either group. We also find that mothers with higher education are more likely to feel stressed and feel that they spend too little time with their child when full-day care is used, this effect is smaller for mothers with lower education. These findings point to greater time constraints faced by mothers whose children are in full-day care compared to half-day care, potentially reducing the capacity to be involved in parenting activities. This effect seems most pronounced for mothers with higher educational attainment. The last three rows of Panel B look at child outcomes.

²²Focusing on one wave and only parents who use center-based care in the time-use data means the sample in the time-use survey is too small to focus on specific parenting activities.

²³As in the time-use data, the household education level is defined by the mother's or single parent's formal education level.

Figure 5.4: Full-day vs half-day care effects by time of day and education, wave 3 only (2012/13)



Notes: Each estimate is based on a separate regression of the outcome summed over center hours or non-center hours on an indicator for usage of full-day center-based care (> 30 vs. 10-30 hours per week) and controls (as in Table 5.2) using the sample of center-based care users. Circles denote lower education households, and squares denote higher education. The filled and hollow shapes together indicate the identified set. See Figure 5.3 for further notes. Appendix Table 5.A.9 reports coefficients along with means of the outcome variables, and the δ required for zero coefficient. Source: German Time-Use Survey (2012/13)

We see evidence for a reduction in children’s nightly sleep. Looking at two measures of children’s well-being, as reported by parents, we find that irritability is not affected whereas perceived happiness of children of lower educated households is somewhat

Table 5.3: The effect of full-day care on parenting and non-parenting activities

	Households with lower maternal education			Households with higher maternal education		
	All parents (1)	Mothers (2)	Fathers (3)	All parents (4)	Mothers (5)	Fathers (6)
<i>Panel A: Parenting activities</i>						
Educational activities						
Reading books or telling stories (daily)	-0.050*** (0.019)	-0.040* (0.023)	-0.073** (0.031)	0.001 (0.015)	0.005 (0.016)	-0.011 (0.026)
Singing or playing instruments (daily)	-0.004 (0.017)	0.019 (0.024)	-0.035 (0.022)	-0.016 (0.017)	-0.013 (0.024)	-0.023 (0.021)
Painting, building or drawing (daily)	-0.018 (0.017)	-0.029 (0.024)	0.003 (0.023)	-0.043*** (0.016)	-0.082*** (0.023)	0.003 (0.022)
Playing games together (daily)	-0.095*** (0.019)	-0.082*** (0.024)	-0.116*** (0.030)	-0.055*** (0.017)	-0.121*** (0.023)	0.021 (0.026)
Recreational activities						
Outdoor activities (daily)	-0.069*** (0.019)	-0.093*** (0.024)	-0.026 (0.029)	-0.054*** (0.017)	-0.088*** (0.024)	-0.012 (0.025)
Gymnastics, sports (daily)	-0.011 (0.019)	0.004 (0.025)	-0.029 (0.031)	-0.009 (0.018)	0.008 (0.024)	-0.035 (0.027)
Watching television or videos (daily)	-0.030 (0.020)	-0.043* (0.025)	-0.010 (0.032)	0.009 (0.019)	-0.001 (0.025)	0.024 (0.027)
<i>Panel B: Non-parenting activities and other outcomes</i>						
Parental outcomes						
Paid work (at least 10 h/w)	0.115*** (0.016)	0.170*** (0.024)	0.027 (0.017)	0.105*** (0.014)	0.185*** (0.023)	0.000 (0.013)
Weekly hours in paid work	3.821*** (0.582)	6.123*** (0.754)	0.451 (0.914)	4.384*** (0.547)	8.232*** (0.752)	-0.244 (0.786)
Personal monthly net income	164.487*** (33.305)	261.447*** (37.959)	24.083 (58.274)	227.728*** (60.269)	419.873*** (51.238)	48.345 (126.902)
Too little time with child (0/1)	0.066*** (0.024)	0.092*** (0.029)	0.026 (0.042)	0.082*** (0.022)	0.186*** (0.028)	-0.050 (0.033)
Feeling stressed (1-5)	0.007 (0.047)	0.057 (0.060)	-0.088 (0.080)	0.145*** (0.041)	0.264*** (0.052)	-0.005 (0.065)
Hours of sleep (parent)	-0.008 (0.047)	-0.002 (0.061)	-0.041 (0.075)	0.001 (0.035)	-0.049 (0.050)	0.049 (0.050)
Child outcomes						
Hours of sleep (child)	-0.174*** (0.051)	-0.193*** (0.067)	-0.131* (0.078)	-0.129*** (0.041)	-0.153*** (0.054)	-0.110* (0.064)
Child is happy and content (1-5)	-0.071** (0.028)	-0.094** (0.037)	-0.032 (0.043)	0.041* (0.023)	0.018 (0.031)	0.064* (0.034)
Child is irritable and cries often (1-5)	-0.032 (0.044)	-0.035 (0.059)	-0.026 (0.067)	-0.033 (0.039)	-0.005 (0.052)	-0.074 (0.058)
Observations	2864	1764	1100	3137	1725	1412

Notes: Table shows conditional coefficients from OLS regressions of the outcome variables on an indicator variable for full-day care (defined as attending center-based care in the morning *and* afternoon) for the sample of center-based care users. Additional controls: dummies for child age, number of children in family, parent and child gender, age of parent, indicator for migrant status, single parent indicator, and an indicator for higher secondary schooling track (for the sample of mothers this drops out due to the split). See Appendix Tables 5.A.10 and 5.A.11 for unconditional coefficients and Oster-bounds. Appendix Table 5.A.12 shows coefficients for an alternative full-day assignment (by hours of usage). Source: pairfam survey 2013-2019.

reduced. This is in line with evidence of negative effects on socio-emotional behavior of full-day care on disadvantaged children (Felfe and Zierow, 2018; Loeb et al., 2007).

5.6 Conclusion

This paper asks: how do parenting activities respond to the usage of center-based care? We outline a framework of potential mechanisms that involve direct effects occurring during center hours and indirect effects outside of those hours. Overall, our analysis shows that using center-based care results in relatively small decreases in parenting activities, which are limited to center hours when time with the child is less activity-rich.

Our analysis of non-parenting activities sheds some light on adjustment mechanisms. In terms of direct effects, the reductions in time with child during center hours tend to coincide with increases in paid work, and reductions in housework and leisure for mothers of both education levels. Thus, center based care is used to take up paid work instead of multi-tasking child care with other activities. Lower education households see a greater reduction in housework and leisure, which helps explain why they have smaller reductions in parenting activities since it suggests that center-based care replaces time with child that is not as activity-rich as it is for higher-education parents. In terms of indirect effects, decreases in sleep and time spent on housework allow lower education parents to maintain a high concentration of activities with children outside of care hours, consistent with either motivation factors or time constraints as explanations.

A specific analysis of the full-day vs. half-day margin finds that using center-based care for 31 hours or more is associated with an additional decrease of parenting activities, with little adjustments during non-center hours. Analyses using survey data shows small reductions in the frequency of certain activities (e.g. 5-12 pp reduction of daily playing) as a result of using full-day care over half-day care, as well as increases in parental stress and some evidence for reduced happiness for children from households with lower maternal education.

Our findings imply a need for greater awareness that development effects of center-based care may come through changes in the home environment not just through the usage of center-based care per se or through quality of this care. Thus, policymakers may want to consider strengthening the home environment channel through the following three measures: (1) Allowing center-based care to ease parental time constraints. Our analysis covers a period when usage of center-based care was expected to facilitate paid employment, and, in the earlier years, this was even the condition for a place. While such conditions may increase the employment effects of center-based care, they may do so at the expense of child development by shutting out one of the mechanisms, i.e. the easing of parental time constraints. (2) Placing a policy focus on the interaction between parents and child care center staff. Care center teachers can help to advise parents with regards to their child's specific developmental progress and

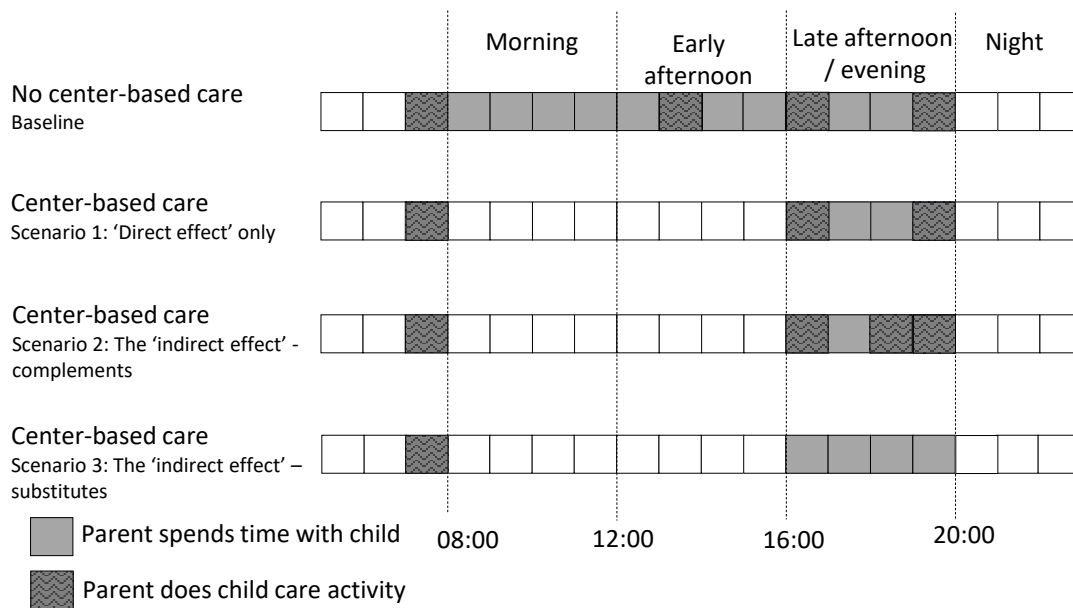
challenges. This may be strengthened by ensuring that teachers have adequate time for interaction with parents. The data in Appendix Table 5.A.6 reveal that most parents have either never sought advice from care center teachers or have done so just once or twice, despite 84 percent of parents reporting a high desire to exchange information about the child (see also Camehl et al., 2015). (3) Improving the quality of center-based care. While our findings highlight an alternative channel for child development effects of center-based care, they should not detract from the importance of child care center quality as one policy priority. Instead, they suggest a complementary way of achieving similar policy goals. There is evidence that qualitatively good center-based care can have positive effects on the quality of the home environment (Kuger et al., 2019).

5.A Appendix

5.A.1 Stylized examples of adjustment mechanisms

Appendix Figure 5.A.1 presents some stylized examples to illustrate various adjustment effects discussed in section 5.3. For simplicity of exposition, we focus the illustration on weekdays and waking hours (7:00 until to 20:00). Effects are illustrated by comparing the ‘no center-based care’ timeline (i.e. the baseline) to the other timeline where center-based care is used. In the ‘no center-based care’ baseline, the parent spends 13 hours with the child, and four of these are spent on parenting activities throughout the day. In scenario 1, the child attends center-based care from 08:00 until 16:00. As a result the child is no longer present with the parent during these hours.²⁴ The direct effect is a decrease in parenting activities in absolute terms of one hour. As parenting activities outside day care hours are unchanged, there is no indirect effect.

Figure 5.A.1: Adjustment of parenting activities with use of center-based care



Notes: Figure illustrates adjustments of time with the child and of parenting activities when center-based care is being used under different scenarios. The upper line shows time use when no center-based care is being used, the bottom three lines show different scenarios when the child is in center-based care. See text for additional details.

²⁴In this simplified framework, we assume a direct relationship between usage of center-based care and time spent with the child. As discussed though, in reality the relationship may be less strong, e.g. in cases where center-based care displaces informal care, e.g. by grandparents.

Scenarios 2 and 3 illustrate the indirect effect, i.e. changes outside of center hours. If center-based care is a complement for parenting activities (scenario 2), it results in an increase of parenting activities in the evening period by one hour in absolute terms. Scenario 3 shows the indirect effect in the substitute case, where there is a reduction by 2 hours in the evening. The overall absolute effect then is a reduction of parenting activities by 3 hours.

Another aspect not covered by the examples—besides night and weekend adjustments—is that center-based care may affect the *type of parenting activities*: Parents might change the share of specific types of parenting activities that are most greatly associated with child development (e.g. reading to the child, see Kalb and van Ours, 2014; Price and Kalil, 2019) This change could work in ways similar to the previous two effects. The usage of center-based care may displace parenting activities of a certain type from one period of the day to another (e.g. if reading is done before sleep rather than during the day). Likewise, usage of center-based care may result in positive or negative indirect effects on particular activities.

5.A.2 Data

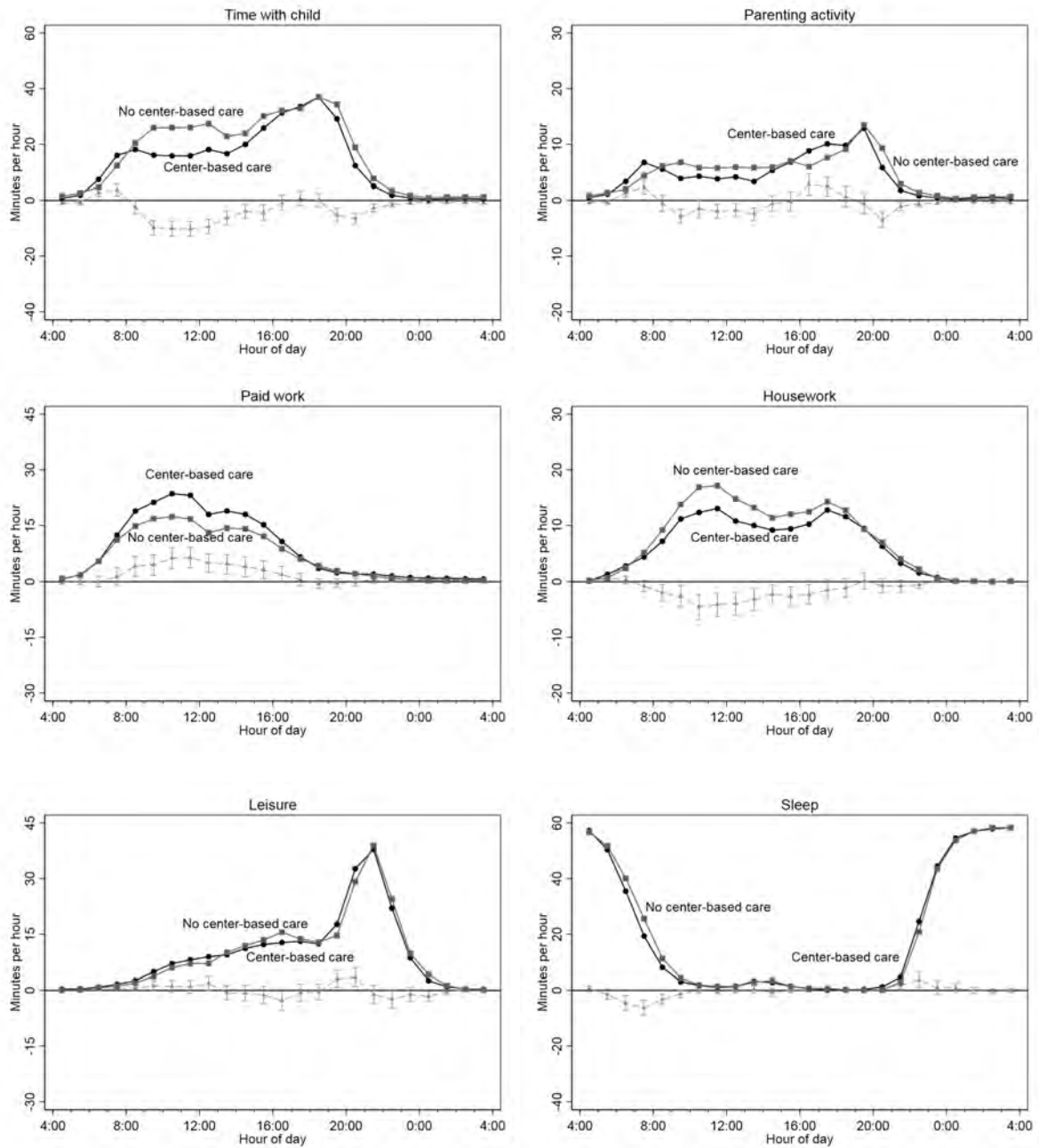
5.A.2.1 Additional descriptive material for time-use data and pairfam

Table 5.A.1: Characteristics of pairfam sample

Variable	Amount of center-based care		Difference
	(1) Half-day	(2) Full-day	
<i>Parental characteristics</i>			
Female (0/1)	0.58	0.58	-0.005 (0.013)
Age in years	35.72	36.15	0.422*** (0.135)
Migration background (0/1)	0.22	0.19	-0.037*** (0.010)
Higher educated (0/1)	0.46	0.53	0.076*** (0.013)
Married (0/1)	0.81	0.70	-0.106*** (0.011)
Paid work (at least 10 h/w, 0/1)	0.71	0.83	0.117*** (0.011)
Weekly hours in paid work	25.66	30.81	5.151*** (0.477)
Personal net income (in Euro)	1426.15	1602.27	176.113*** (40.560)
Household net income (in Euro)	3538.35	3638.63	100.281* (59.046)
<i>Child characteristics</i>			
Girl (0/1)	0.49	0.50	0.013 (0.013)
Age in years	4.52	4.57	0.052** (0.022)
Number of siblings	1.43	1.28	-0.153*** (0.025)
Observations	3345	2660	6005

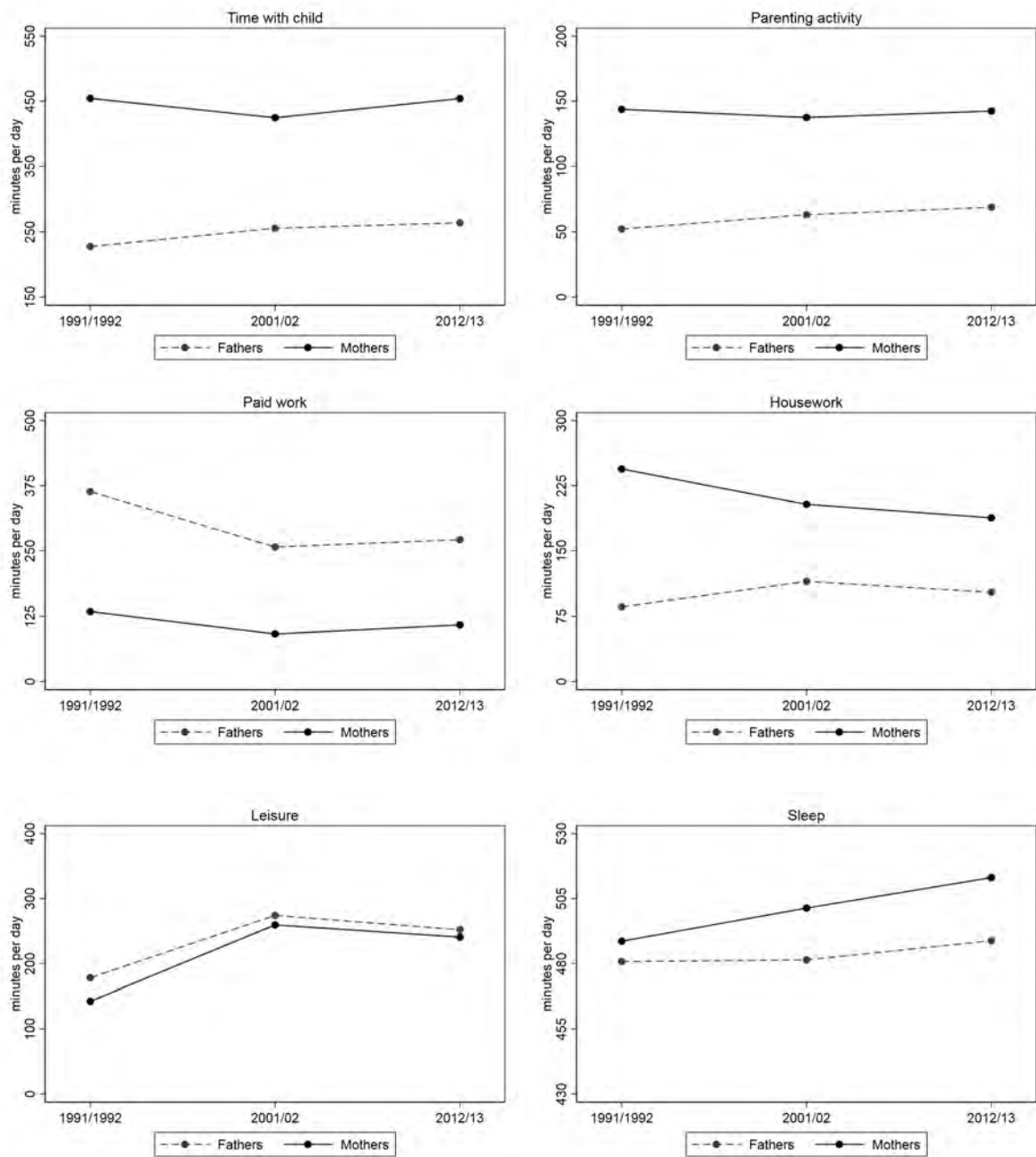
Notes: Full-day child care indicates usage of center-based care in the morning *and* afternoon. Half-day care morning *or* afternoon. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: pairfam, 2013-2019

Figure 5.A.2: Parents' activities by usage of center-based care



Notes: Circles denote parents with a child in center-based care, squares those without. Differences and averages are estimated in weighted regressions with indicators for child age and evaluated at mean values. Whiskers indicate 95% confidence intervals. Data consists of time slots in ten minute intervals (five in the first survey wave) which then are aggregated by hour of day. Sample includes weekdays (68%) and weekend days (32%), pools mothers and fathers. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

Figure 5.A.3: Average time use for mothers and fathers by survey wave



Notes: Coefficients are obtained by regressing activities on an indicator for mothers (vs. fathers) with child-age indicators and then evaluating means at average values (regressions are weighted). Sample consists of weekdays and weekend days. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

5.A.2.2 Activities categories in time-use data

Table 5.A.2: Overview of activities in time-use data, 2012/13 wave

Broad activity (1-digit)	German title of 1-digit activity	# of 3-digit activities	Examples of 3-digit activities
Personal care	"Persönlicher Bereich / Physiologische Regeneration"	5	Sleep, eating and drinking, washing and dressing, ...
Paid work	"Erwerbstätigkeit"	9	Main work, secondary work, On-the job training, ...
Qualifications / Education	"Qualifikation / Bildung"	29	German lessons, higher education, training outside of work hours, ...
Household and family care	"Haushaltsführung und Betreuung der Familie"	43	Preparing meals, shopping, small repairs, ...
Voluntary work	"Ehrenamtliche Tätigkeit / Freiwilligenarbeit / Unterstützung für andere Haushalte / Teilnahme an Versammlungen"	5	Voluntary work, supporting other households, political events, ...
Social life and entertainment	"Soziales Leben und Unterhaltung"	14	Talking (with friends), cinema, relaxation, ...
Sport, hobbies and games	"Sport / Hobbys / Spiele"	20	Going for a walk, hunting / fishing, computer games, ...
Media usage	"Mediennutzung"	13	Reading newspaper, watching TV, communication with computer or smartphone, ...
Travel time	"Zweckbestimmte Wegezeiten und Hilfscodes"	27	Travel time to main work, travel time to school, travel time to visit friends, ...

Notes: Table summarizes the broad (1-digit) activities that are reported in the German time-use data set. The English-language activity labels are our own translation from the tables available with the time-use survey data for 2012/2013. Full tables for each wave (in German) can be accessed at the website from the research data center of the German Federal Statistical Office:

<https://www.forschungsdatenzentrum.de/de/haushalte/zve>

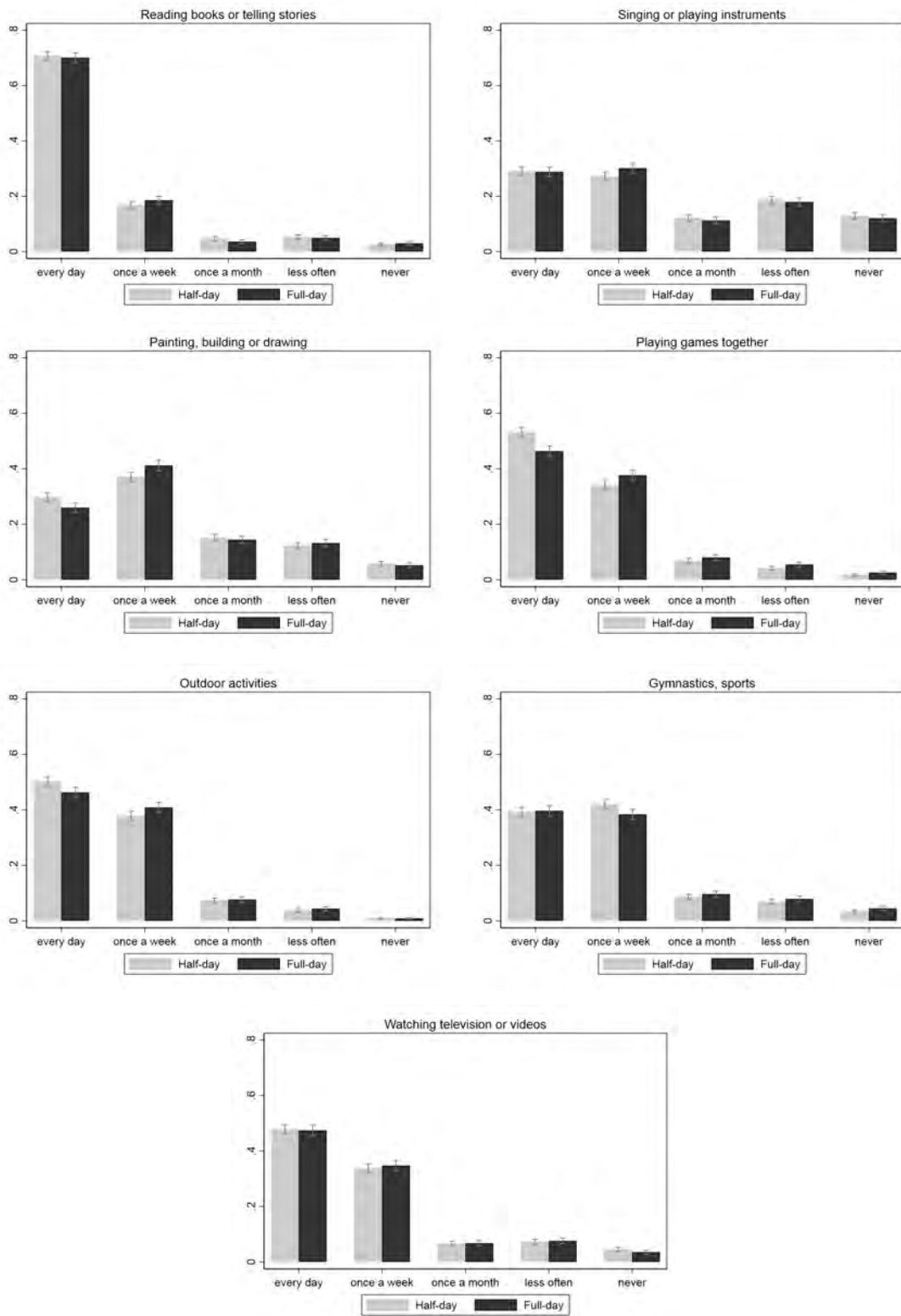
Table 5.A.3: List of detailed parenting activities in time-use data, 2012/13 wave

Code	Activity	German (original)
<i>2-digit category</i>		
47	Child care	"Kinderbetreuung im Haushalt"
<i>3-digit category</i>		
471	Primary care, hygiene and supervision	"Körperpflege und Beaufsichtigung"
472	Assisting homework / giving instructions to child	"Hausaufgabenbetreuung/Anleitungen geben"
473	Playing and doing sports with child	"Spielen und Sport mit Kindern"
474	Talking with child	"Gespräche mit Kindern im Haushalt"
475	Accompanying child / realising appointments with child	"Kind begleiten/Termine mit dem Kind wahrnehmen"
476	Reading to child / telling stories	"Kindern vorlesen/Geschichten erzählen"
479	Other activities with child	"Sonstige Aktivitäten im Bereich Kinderbetreuung"

Notes: Table reports the detailed (3-digit) parenting activities reported in the time-use data set, 2012/13 wave. The English-language activity labels are our own translation from the tables available with the time-use survey data for 2012/2013. Full tables for each wave (in German) can be accessed at website for the research data center of the German Federal Statistical Office:

<https://www.forschungsdatenzentrum.de/de/haushalte/zve>

Figure 5.A.4: Shared activities with the child by half- or full-day usage of center-based care



Notes: Figure shows the frequency of activities of mothers or fathers with their children (in the previous three months). Whiskers show 95% confidence intervals. Source: pairfam, 2013-2019.

5.A.2.3 Data on informal care

Table 5.A.4: Weekly hours in care - SOEP

Variable	Obs	Mean	Std. Dev.	P25	P50	P75
<i>All care types</i>	14311	21.447	20.145	1	20	37
<i>Informal care (outside the household)</i>	14311	5.055	9.433	0	2	6
>0 hours (0/1)	14311	.554	.497	0	1	1
>20 hours (0/1)	14311	.05	.217	0	0	0
>30 hours (0/1)	14311	.02	.141	0	0	0
Family	14311	4.622	8.943	0	1	6
Other informal	14311	.433	3.114	0	0	0
<i>Center-based care</i>	14311	16.392	17.28	0	15	30
>0 hours (0/1)	14311	.52	.5	0	1	1
>20 hours (0/1)	14311	.416	.493	0	0	1
>30 hours (0/1)	14311	.243	.429	0	0	0
Center-based care	14311	15.614	16.846	0	0	30
Center-based care (conditional on usage)	7218	31.325	8.784	25	30	40
<i>Age of child (in months)</i>	14311	33.588	23.072	12	31	63

Notes: Sample consists of children aged 0-72 months. Averages are calculated using survey weights. All care types include all forms of care indicated besides care provided by the respondent or the partner. Family care consists of care by the partner (if not living in the household), grandparents, older siblings and other relatives. Other informal care arrangements are nannies or a residual *other* category. Formal care reflects hours spent at either center-based care (95.1% in our data) or with publicly funded family day care (4.9%). Sample covers survey years 2010-2018. Data from the German Socio-Economic Panel (SOEP v35), which is a long-running household survey containing information on about 15,000 households per year (Goebel et al., 2019).

Table 5.A.5: Usage of formal and informal care

	Below 3		Above 3		Below 3		Above 3		All
	Center-based care				Full-day care				
	No	Yes	No	Yes	No	Yes	No	Yes	
Weekly hours at center-based care	0.00 (0.00)	28.56 (12.06)	0.00 (0.00)	28.80 (11.01)	23.60 (11.27)	33.71 (10.62)	24.93 (9.76)	33.61 (10.57)	21.26 (15.91)
Family care in morning	0.18 (0.39)	0.03 (0.18)	0.14 (0.35)	0.03 (0.16)	0.05 (0.21)	0.02 (0.15)	0.03 (0.16)	0.03 (0.16)	0.07 (0.25)
Family care in afternoon	0.23 (0.42)	0.25 (0.43)	0.21 (0.41)	0.28 (0.45)	0.28 (0.45)	0.22 (0.42)	0.31 (0.46)	0.24 (0.43)	0.26 (0.44)
Family care - any time	0.25 (0.44)	0.26 (0.44)	0.24 (0.43)	0.28 (0.45)	0.29 (0.45)	0.22 (0.42)	0.31 (0.46)	0.24 (0.43)	0.27 (0.44)
Other informal care in morning	0.02 (0.12)	0.00 (0.05)	0.03 (0.16)	0.00 (0.05)	0.00 (0.05)	0.00 (0.06)	0.00 (0.05)	0.00 (0.04)	0.01 (0.08)
Other informal care in afternoon	0.02 (0.14)	0.03 (0.17)	0.04 (0.20)	0.04 (0.20)	0.03 (0.18)	0.03 (0.16)	0.04 (0.20)	0.04 (0.19)	0.03 (0.18)
Other informal care - any time	0.02 (0.14)	0.03 (0.17)	0.04 (0.20)	0.04 (0.20)	0.03 (0.18)	0.03 (0.16)	0.04 (0.20)	0.04 (0.19)	0.03 (0.18)
Observations	2560	1871	226	5991	963	908	3336	2655	10648

Notes: Sample consists of children aged 0-72 months. Columns are split by age of the child (0-2 vs. 3-5 years) and by usage of center-based care. Full-day care is defined as using center-based care in the morning and afternoon in contrast to only one of these (thus conditional on day care usage). Family care includes grandparents, siblings and other relatives. Other informal care arrangements consist of friends, a nanny in-house, and other non-relatives. Source: pairfam, 2013-2019.

5.A.3 Center-based care activities with children

This appendix section examines the activities children are exposed to at center-based care.²⁵ We use two data sets for this; first, the National Educational Panel Study *NEPS*, in which the Starting Cohort Kindergarten (SC2) contains interviews with educators and heads of child care centers (Blossfeld and von Maurice, 2011). This allows for looking at the regularity of various activities performed at the institutions. As a second data source we use the data set *Educational Processes, Competence Development and Selection Decisions in Preschool and School Age (BiKS-3-10)*, which started in two German states (Hesse and Bavaria) in 2005. The starting sample of BiKS consisted of 550 children from 97 child care centers (Weinert et al., 2013). Educators were (among other aspects) asked about the regularity and duration of extracurricular activities as well as the broader institutional environment. Parents were further asked detailed questions about their children and their assessment of the child care centers.

In Panel A of Appendix Table 5.A.6 the frequency of regular activities are shown. The activities listed are all arguably enhancing cognitive development (e.g. books, puzzles, number games, musical activities) or motor skills (e.g. tinkering, sports). Although no information on the minutes per activity are included in the data, it is evident that the educational content is relatively high, as many activities are being performed daily or even several times per day. Panel B displays the frequency of extracurricular activities and — conditional on offerings — the average length of these. Most institutions offer extra activities, usually once or twice a week. Although these findings are not nationally representative, as the BiKS-3-10 data stems from two West German states, it suggests that children are not merely being supervised at center-based care but that they are often exposed educational activities.

²⁵There are some obvious caveats to this; we do not know how often and how long children take part in activities if they are performed at the group level and they will less frequently experience one-to-one interactions in center-based care (Clarke-Stewart et al., 1994). Many activities can also be less beneficial for children if they are conducted in groups rather than in one-on-one interactions (thus perhaps requiring more exposure time at center-based care compared to at home).

Table 5.A.6: Activities in center-based care

	(1)	(2)	(3)	(4)
<i>Panel A: NEPS SC2</i>				
Frequency of regular activities (share)	Several times per day	Daily	Several times per week	Weekly
Books / letter games	0.445	0.086	0.102	0.009
Puzzles	0.515	0.065	0.067	0.003
Number games	0.408	0.089	0.127	0.015
Building things / tinkering	0.581	0.046	0.027	0.001
Musical activities	0.195	0.241	0.173	0.039
Sports	0.203	0.160	0.202	0.082
Experiencing nature	0.091	0.134	0.183	0.071
Observations	2775	2775	2775	2775
<i>Panel B: BiKS-3-10</i>				
Extra curriculum activities	Offered (share)	Weekly frequency	Minutes per offering	Minutes per week
Any activity	0.919			
Sport	0.760	1.205	29.338	32.891
Foreign languages	0.349	1.377	11.446	9.422
Craft activities	0.327	1.688	11.774	17.008
Nature studies	0.524	1.015	36.868	24.610
School preparation	0.837	2.018	60.583	79.240
Musical activities	0.645	1.705	16.100	19.800
Observations	172	172	172	172
<i>Panel C: BiKS-3-10</i>				
Parental responses to center-based care attendance				
Center-based care attendance enriched relationship with child	No	Yes		
	0.297	0.703		
Have sought advice for child rearing by care center staff	Never	Once or twice	Several times	
	0.334	0.417	0.248	
Desire to exchange information about child	Unwilling	Rather high	High	
	0.025	0.139	0.836	
Observations		438	438	438

Notes: Panel A shows the frequency of regular group activities in child care centers. Activities are coded on a seven point scale from less than once a month to several times a day. Panel B shows extra curriculum activities offered at child care centers. Columns (2) and (3) in Panel B are conditional on offerings. Panel C show parental reactions to care center enrollment of their child. Source: NEPS SC2 (2011-2012) and BiKS-3-10 (2005-2006).

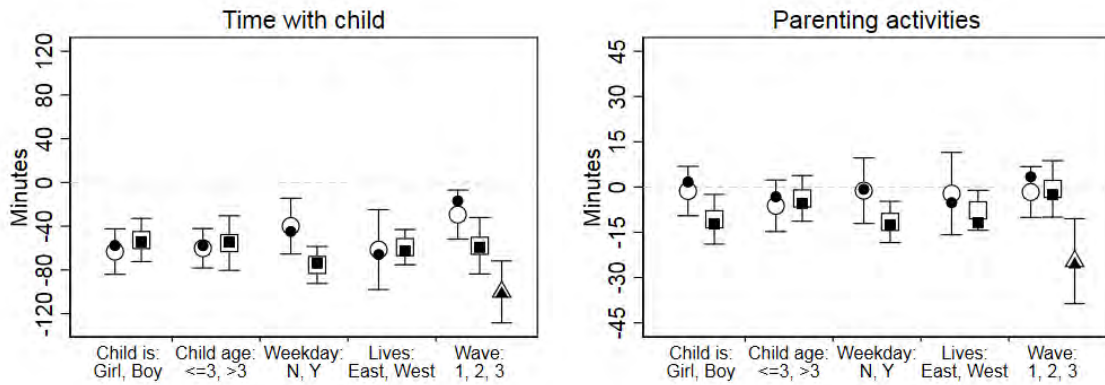
5.A.4 Heterogeneous effects

We explore further heterogeneities of the effect of center-based care on parenting activities. In Appendix Figure 5.A.5, we split the sample by sex of the child, (male/female), by child age (under and over three years), by day of the week (weekdays, weekend days) by location (East/West Germany), and by survey wave (1991/92, 2001/02, 2012/13). Some of these sample splits are motivated by the different center-based care environments for different age groups, different regions and over time (see section 5.2): in East Germany enrollment rates have always been substantially higher and, since the mid-2000s, the whole of Germany has seen a strong increase in enrollment for under threes and in full-day care for all age groups (Jessen et al., 2018).

For both boys and girls, using center-based care reduces parents' time spent with the child but for boys there is a nearly proportionate decrease in parenting activities whereas, for girls, parents continue to maintain the same absolute level of child care. In Appendix Figure 5.A.6, we investigate this further, finding that the effect for girls is driven by positive indirect effects. In particular, there is a significant increase in parenting activities by mothers outside of center hours despite a decrease in time spent with the child in those hours. Overall, same-sex interactions appear to be positively effected by center-based care usage. This result may be explained by, e.g., research from the U.S. that shows mothers spend more time on activities with daughters and fathers spend more time on activities with sons (Lundberg et al., 2007).

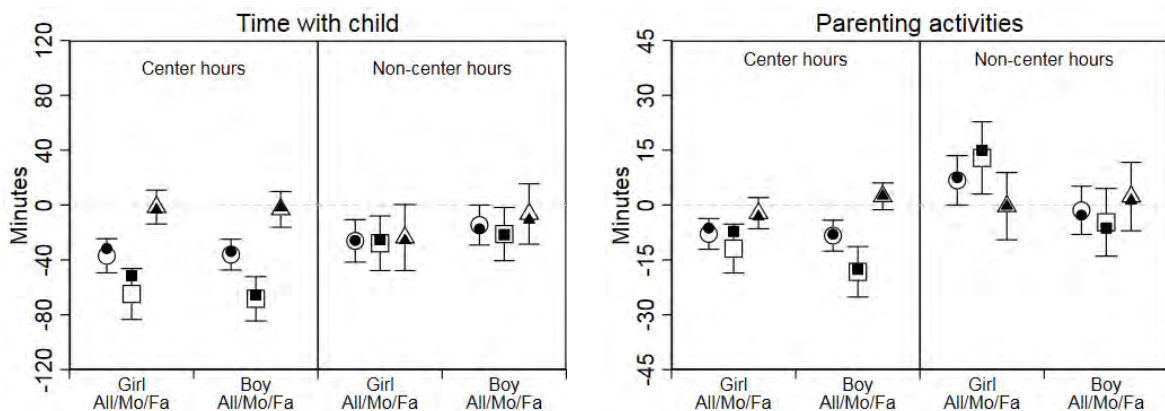
In Appendix Figure 5.A.5 there is little heterogeneity by child age, nor by region. By survey wave, we observe an increased magnitude of effects for later waves, consistent with more child care center places and longer average hours of care in more recent years.

Figure 5.A.5: Heterogeneity in overall effects on parenting activities – mothers and fathers pooled



Notes: Plots show heterogeneities in effects of center-based care on parenting activities. Circles denote the respective first, squares the second and triangles (if applicable) the third group. Estimates are based on separate sub-sample regressions of the outcome variable on a center-based care indicator and controls (see notes to Table 5.2 for details). Waves 1, 2, and 3 correspond to the time-use survey waves 1991/92, 2001/02, and 2012/13 respectively. The hollow shapes and whiskers indicate conditional coefficient ($\delta = 0$) and the 90% confidence intervals. The filled shapes indicate estimates under the assumption of $\delta = 1$, i.e. equally large selection on unobservables as on observables. The filled and hollow shapes indicate the identified set. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

Figure 5.A.6: Effects of center-based care on parenting activities for boys and girls

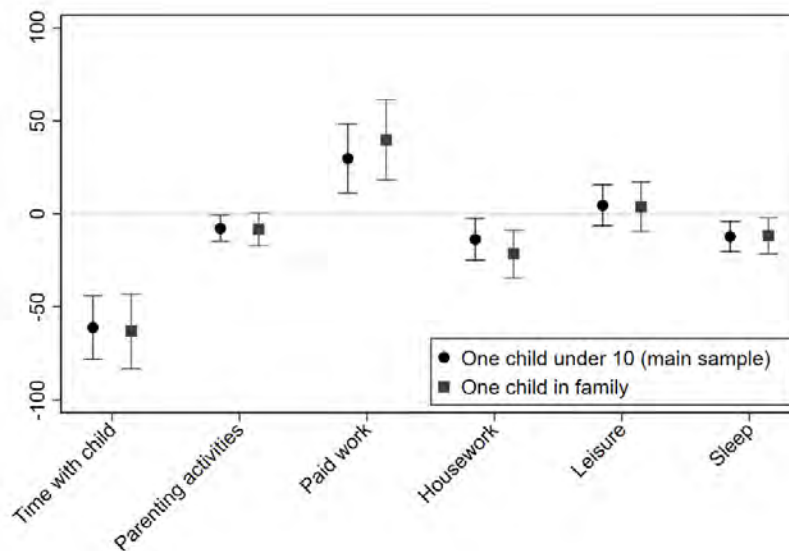


Notes: Plots show heterogeneities by gender of child in effects of center-based care on parenting activities. Circles denote mothers and fathers pooled, squares denote mothers and triangles fathers. See Figure 5.3 for further notes. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

5.A.5 Tests of sample restrictions

In this appendix section, we compare coefficients when different sample restrictions are imposed. Our main analysis sample with the time-use data is restricted to families with one child *under 10 years*. In Appendix Figure 5.A.7, we compare coefficients when we tighten the requirement and impose that only one child *of any age* is in the family (this reduces the observation number from 4,295 to 2,984). The reason for this is that although we know that the outcome *time with child* is constructed in the survey such that it only refers to children under 10, other parenting activities could still be conducted with older children (although these are arguably mostly performed with younger children and not with those of secondary school age). Coefficients in Appendix Figure 5.A.7 from both samples are remarkably similar and statistically indistinguishable.

Figure 5.A.7: Comparison of coefficients by sample restriction I

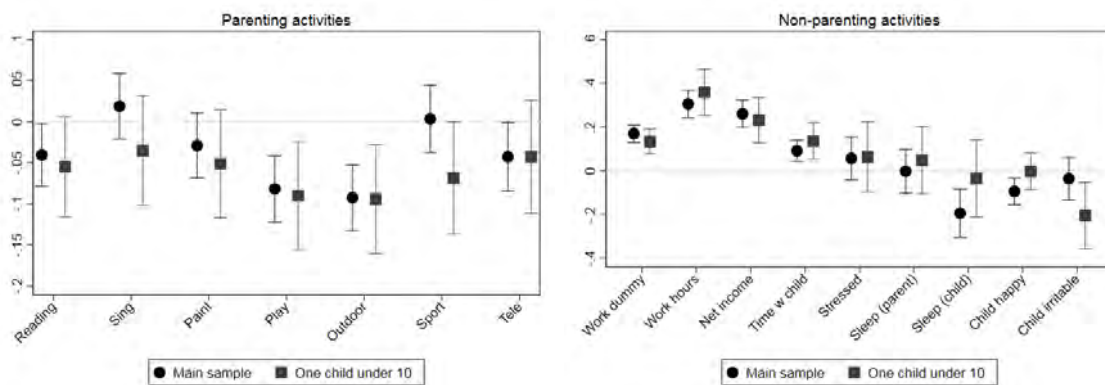


Notes: Figure shows coefficients and 95% confidence intervals for the main analysis sample (one child under 10 years) and for a tighter sample restriction of one child of *any age* in families. Estimates refer to mothers and fathers pooled, and concern the whole day. Coefficients based on conditional specification with control variables as indicated in Table 5.2. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

In Appendix Figure 5.A.8, we investigate to what degree the data driven sample restriction in the time-use data of one child under 10 years reduces the external validity of the findings, i.e. would the findings also hold for households with more children under 10? The household survey (pairfam) does not require the same sample restriction as the time-use survey as questions are child-specific (but it contains the information needed to impose the same sample restriction). Thus we compare the coefficients shown in

Table 5.3 obtained using the unrestricted sample (i.e. with potentially several children in this age group in one household) and apply the same restriction that we use in the time-use data. Appendix Figure 5.A.8 shows that, for parenting activities (left panel), coefficients are quite similar and all confidence intervals overlap. For non-parenting activities and other outcomes (right panel) coefficients are again comparable. Overall this suggests that the sample restriction imposed do not severely threaten the generalizability of the findings.

Figure 5.A.8: Comparison of coefficients by sample restriction II



Notes: Figure shows coefficients and 95% confidence intervals for the unrestricted sample (*main sample*) and when applying the same sample restriction as in the time-use data (*one child under 10*). Estimates refer to mothers, i.e. the main sample estimates correspond to column (2) of Table 5.3. For presentation purposes coefficient and confidence intervals for working hours and net income are rescaled by a factor of 20 and 1000, respectively. $N = 6,005$ for the main sample and $N = 1,866$ for the one child under 10 sample. Source: pairfam, 2013-2019.

5.A.6 Further result tables

Table 5.A.7: Effects of center-based care on parents' time spent on parenting and non-parenting activities, by time of day and education

	Households with lower maternal education			Households with higher maternal education		
	Center hrs (1)	Evening and weekend (2)	Night (3)	Center hrs (4)	Evening and weekend (5)	Night (6)
<i>Outcome: Time with child (in minutes)</i>						
Unconditional	-54.5*** (7.67)	-15.7 (10)	-16*** (4.24)	-37.3*** (9.63)	-3.41 (12)	-3.77 (4.09)
Conditional	-42.1*** (6.39)	-14.6** (6.93)	-17.6*** (4.21)	-32*** (8.01)	-6.13 (7.82)	-1.76 (4.25)
Mean	81.777	202.865	51.833	87.844	214.340	58.205
Identified set	[-42.123, -37.005]†	[-14.637, -14.245]†	[-18.309, -17.630]†	[-32.000, -29.890]†	[-7.148, -6.133]†	[-1.763, -0.784]†
δ for 0 coefficient	5.784	15.314	88.276	8.653	-8.807	1.721
R^2 (unc., con.)	(0.015, 0.396)	(0.015, 0.574)	(0.015, 0.100)	(0.028, 0.368)	(0.028, 0.591)	(0.028, 0.109)
<i>Outcome: Parenting activities (in minutes)</i>						
Unconditional	-8.2*** (2.28)	3.58 (3.27)	-4.34*** (1.35)	-13.7*** (3.51)	5.51 (4.51)	-4.31** (2.03)
Conditional	-6.16*** (2.14)	3.72 (3.07)	-3.85*** (1.39)	-12.8*** (3.2)	4.3 (4.33)	-3.92* (2.07)
Mean	21.475	48.515	20.530	24.636	57.923	24.128
Identified set	[-6.159, -5.146]†	[3.724, 3.791]†	[-3.854, -3.578]†	[-12.779, -12.316]†	[3.706, 4.304]†	[-3.921, -3.662]†
δ for 0 coefficient	4.700	29.036	7.651	10.116	5.259	7.057
R^2 (unc., con.)	(0.077, 0.295)	(0.077, 0.263)	(0.077, 0.215)	(0.077, 0.299)	(0.077, 0.261)	(0.077, 0.171)
<i>Outcome: Paid work (in minutes)</i>						
Unconditional	41.2*** (10.9)	-3.04 (4.68)	13.5*** (4.87)	21.8* (11.9)	-1.27 (4.64)	-4.4 (3.95)
Conditional	27.1*** (8.83)	-2.71 (4.82)	13.8*** (5.05)	19.8** (10)	-5.51 (4.8)	-1.9 (3.9)
Mean	151.320	35.234	34.718	140.418	32.194	26.917
Identified set	[21.943, 27.143]†	[-0.271, 0.748]	[13.763, 13.849]†	[19.035, 19.799]†	[-0.551, -0.262]†	[-1.900, -0.950]†
δ for 0 coefficient	4.406	0.274	26.517	12.066	1.824	1.912
R^2 (unc., con.)	(0.007, 0.432)	(0.007, 0.088)	(0.007, 0.160)	(0.007, 0.391)	(0.007, 0.072)	(0.007, 0.131)
<i>Outcome: Housework (in minutes)</i>						
Unconditional	-17.7*** (5.74)	3.07 (5.14)	-1.67 (1.81)	-24.4*** (5.88)	-9.37 (5.81)	-1.04 (1.94)
Conditional	-10.9** (4.75)	-1.07 (4.45)	-2.09 (1.89)	-15.9*** (5.22)	-3.28 (4.98)	-0.766 (1.96)
Mean	61.557	87.069	23.618	55.077	81.546	22.814
Identified set	[-10.893, -8.424]†	[-2.537, -1.075]†	[-2.245, -2.094]†	[-15.915, -12.766]†	[-3.280, -1.066]†	[-0.777, 0.289]
δ for 0 coefficient	3.855	-0.788	-33.059	4.188	1.451	0.217
R^2 (unc., con.)	(0.001, 0.364)	(0.001, 0.379)	(0.001, 0.130)	(0.003, 0.303)	(0.003, 0.344)	(0.003, 0.062)
<i>Outcome: Leisure (in minutes)</i>						
Unconditional	-8.08*** (3.01)	-3.29 (5.53)	-2.28 (4.13)	7.97** (3.51)	11.3* (6.22)	2.45 (4.19)
Conditional	-7.79*** (2.93)	.441 (4.64)	3.44 (4.26)	4.71 (3.17)	7.52 (5.44)	1.25 (4.27)
Mean	23.670	77.361	106.861	28.005	83.611	104.325
Identified set	[-7.791, -7.683]†	[0.441, 1.766]†	[3.440, 5.514]†	[3.490, 4.712]†	[6.149, 7.520]†	[0.736, 1.247]†
δ for 0 coefficient	17.399	-0.350	-1.791	3.395	4.568	2.272
R^2 (unc., con.)	(0.003, 0.183)	(0.003, 0.378)	(0.003, 0.063)	(0.014, 0.213)	(0.014, 0.369)	(0.014, 0.093)
<i>Outcome: Sleep (in minutes)</i>						
Unconditional	.8 (2.35)	-1.85 (2.52)	-10.4* (5.34)	-3.44 (2.63)	-3.1 (2.22)	3.47 (5.32)
Conditional	.948 (2.55)	-1.08 (2.24)	-15.6*** (5.51)	-3.05 (2.43)	-2.69 (2.07)	.392 (5.37)
Mean	10.705	15.711	463.680	9.335	13.755	464.920
Identified set	[0.948, 1.010]†	[-1.081, -0.802]†	[-17.643, -15.647]†	[-3.054, -2.896]†	[-2.693, -2.544]†	[-0.847, 0.392]
δ for 0 coefficient	-60.860	3.475	-10.858	9.713	10.287	0.327
R^2 (unc., con.)	(0.008, 0.049)	(0.008, 0.253)	(0.008, 0.100)	(0.010, 0.053)	(0.010, 0.255)	(0.010, 0.096)
Observations	2482	2482	2482	2008	2008	2008

Notes: Center hours are from 8am-4pm on weekdays, evening and weekend consists of 4pm-8pm on weekdays and entire weekend days (8am-8pm). Nights are from 8pm-8am. Table shows coefficients from OLS regressions of the outcome variables on an indicator variable for center-based care usage. Figure 5.3 shows the conditional coefficients and the coefficient under the assumption of equally large selection on observables as on unobservables ($\delta = 1$). See Table 5.2 for other table notes and section 5.4 for details on the empirical specification. Source: German Time-Use Survey (1991/92, 2001/02 and 2012/13)

Table 5.A.8: Effects of center-based care on parents' time spent on specific parenting activities, by time of day and education

	Households with lower educated mothers			Households with higher educated mothers		
	Center hrs (1)	Evening and weekend (2)	Night (3)	Center hrs (4)	Evening and weekend (5)	Night (6)
<i>Outcome: Reading (in minutes)</i>						
Unconditional	-.0188 (.201)	1.61** (.739)	-.661 (.425)	-.114 (.351)	.182 (.603)	.638 (.393)
Conditional	.00923 (.25)	1.33* (.788)	-.72 (.452)	.0102 (.343)	-.347 (.6)	.557 (.427)
Mean	0.286	1.818	0.774	0.538	2.227	1.121
Identified set	[0.009, 0.021]†	[1.138, 1.327]†	[-0.755, -0.720]†	[0.010, 0.074]†	[-0.646, -0.347]†	[0.505, 0.557]†
δ for 0 coefficient	-0.814	4.539	81.341	-0.169	-1.267	6.545
R^2 (unc., con.)	(0.023, 0.026)	(0.023, 0.063)	(0.023, 0.057)	(0.029, 0.037)	(0.029, 0.056)	(0.029, 0.064)
<i>Outcome: Playing (in minutes)</i>						
Unconditional	-3.38* (1.76)	-2.03 (3.77)	.866 (.833)	-5.12** (2.39)	1.42 (4.25)	-2.05* (1.12)
Conditional	-2.88* (1.72)	-2.95 (3.71)	1.22 (.844)	-5.32** (2.33)	.766 (4.15)	-1.54 (1.18)
Mean	6.330	23.224	1.827	7.706	27.242	3.318
Identified set	[-2.877, -2.634]†	[-3.357, -2.946]†	[1.220, 1.430]†	[-5.425, -5.324]†	[0.417, 0.766]†	[-1.540, -1.135]†
δ for 0 coefficient	7.481	-10.885	-6.722	31.510	2.075	3.017
R^2 (unc., con.)	(0.011, 0.108)	(0.011, 0.138)	(0.011, 0.026)	(0.034, 0.136)	(0.034, 0.152)	(0.034, 0.062)
<i>Outcome: Talking (in minutes)</i>						
Unconditional	-.185 (.444)	-.486 (.456)	.0854 (.181)	.215 (.221)	-.773 (.721)	.071 (.274)
Conditional	-.299 (.483)	-1.07** (.504)	-.285 (.218)	.392 (.257)	-.895 (.778)	-.147 (.264)
Mean	0.614	1.263	0.766	0.688	1.009	0.703
Identified set	[-0.343, -0.299]†	[-1.333, -1.072]†	[-0.537, -0.285]†	[0.392, 0.465]†	[-0.957, -0.895]†	[-0.340, -0.147]†
δ for 0 coefficient	-9.326	-3.888	-1.215	-6.709	-40.393	-0.815
R^2 (unc., con.)	(0.024, 0.047)	(0.024, 0.047)	(0.024, 0.048)	(0.014, 0.049)	(0.014, 0.026)	(0.014, 0.023)
<i>Outcome: Primary care (in minutes)</i>						
Unconditional	-4.01* (2.08)	4 (2.69)	-5.12*** (1.67)	-7.23*** (2.56)	5.89** (3)	-1.51 (2.1)
Conditional	-3.36 (2.11)	3.5 (2.6)	-5.95*** (1.65)	-7.65*** (2.4)	5.18* (2.86)	-1.37 (2.08)
Mean	9.175	21.103	15.337	11.405	26.338	18.752
Identified set	[-3.357, -2.928]†	[3.193, 3.500]†	[-6.403, -5.954]†	[-7.923, -7.649]†	[4.738, 5.177]†	[-1.365, -1.256]†
δ for 0 coefficient	5.413	6.990	-44.218	36.166	7.239	7.029
R^2 (unc., con.)	(0.074, 0.276)	(0.074, 0.288)	(0.074, 0.215)	(0.108, 0.291)	(0.108, 0.272)	(0.108, 0.205)
<i>Outcome: Residual child care activities (in minutes)</i>						
Unconditional	.245 (.981)	.28 (1.09)	-.826 (.624)	-2.97* (1.52)	.103 (.897)	-.102 (.465)
Conditional	.535 (1.38)	-.267 (1.17)	-1.07* (.646)	-2.17 (1.56)	-.0956 (1)	.448 (.567)
Mean	4.242	3.838	1.296	4.088	3.490	1.457
Identified set	[0.535, 0.641]†	[-0.539, -0.267]†	[-1.168, -1.066]†	[-2.167, -1.862]†	[-0.198, -0.096]†	[0.448, 0.674]†
δ for 0 coefficient	-6.366	-1.060	-15.918	5.604	-1.026	-2.073
R^2 (unc., con.)	(0.010, 0.049)	(0.010, 0.021)	(0.010, 0.053)	(0.005, 0.085)	(0.005, 0.035)	(0.005, 0.026)
Observations	1188	1188	1188	1338	1338	1338

Notes: Center hours are from 8am-4pm on weekdays, evening and weekend consists of 4pm-8pm on weekdays and entire weekend days (8am-8pm). Nights are from 8pm-8am. Table shows coefficients from OLS regressions of the outcome variables on an indicator variable for center-based care usage. See Table 5.2 for other table notes and section 5.4 for details on the empirical specification. Source: German Time-Use Survey (2001/02 and 2012/13)

Table 5.A.9: Effects of full-day vs. half-day center-based care on parenting and non-parenting activities using one wave (2012/13), by time of day and education

	Households with lower maternal education			Households with higher maternal education		
	Center hrs (1)	Evening and weekend (2)	Night (3)	Center hrs (4)	Evening and weekend (5)	Night (6)
<i>Outcome: Time with child (in minutes)</i>						
Unconditional	-12.2 (10.2)	35.3 (22.1)	18.6*** (6.1)	-37.9*** (10.2)	29.5* (16.9)	9.08* (5.11)
Conditional	-3.87 (10.1)	10.9 (14.4)	14.4** (6.12)	-35.3*** (8.81)	9.53 (13)	13.2** (5.29)
Mean	58.747	195.620	47.009	70.864	204.424	53.464
Identified set	[-3.865, -0.737]†	[1.813, 10.899]†	[11.721, 14.351]†	[-35.305, -34.240]†	[2.541, 9.526]†	[13.202, 14.846]†
δ for 0 coefficient	1.222	1.189	3.654	11.934	1.347	-7.472
R^2 (unc., con.)	(0.075, 0.283)	(0.075, 0.592)	(0.075, 0.183)	(0.014, 0.296)	(0.014, 0.520)	(0.014, 0.093)
<i>Outcome: Parenting activities (in minutes)</i>						
Unconditional	-5.49 (3.55)	-7.22 (6.84)	3.92 (2.93)	-10.4*** (3.45)	1.38 (5.68)	-3.86 (2.4)
Conditional	-5.99* (3.48)	-1.56 (6.98)	4.31 (2.94)	-7.35** (3.04)	-2.25 (5.46)	-2.99 (2.32)
Mean	13.404	41.644	15.933	16.363	51.640	19.225
Identified set	[-6.187, -5.990]†	[-1.918, -1.564]†	[4.307, 4.466]†	[-7.348, -6.028]†	[-4.085, -2.252]†	[-2.991, -2.523]†
δ for 0 coefficient	52.956	-6.090	61.694	4.370	-1.357	4.777
R^2 (unc., con.)	(0.021, 0.208)	(0.021, 0.160)	(0.021, 0.192)	(0.045, 0.230)	(0.045, 0.208)	(0.045, 0.127)
<i>Outcome: Paid work (in minutes)</i>						
Unconditional	-4.93 (20.1)	13.4 (8.31)	-6.92 (7.72)	17.3 (16.3)	-4.88 (6.89)	4.81 (5.6)
Conditional	3.51 (15.7)	7.24 (8.24)	-4.45 (6.31)	35.3*** (13.3)	.441 (6.99)	6.56 (5.6)
Mean	171.796	36.315	38.017	155.624	34.609	28.453
Identified set	[3.512, 6.711]†	[4.674, 7.239]†	[-4.448, -3.189]†	[35.332, 41.596]†	[0.441, 2.595]†	[6.564, 7.205]†
δ for 0 coefficient	-1.214	2.519	3.027	-6.630	-0.213	-15.863
R^2 (unc., con.)	(0.045, 0.458)	(0.045, 0.097)	(0.045, 0.160)	(0.006, 0.405)	(0.006, 0.100)	(0.006, 0.111)
<i>Outcome: Housework (in minutes)</i>						
Unconditional	-7.41 (9.98)	-5.72 (11.6)	5.28* (2.71)	-15.9** (6.65)	7.75 (7.6)	-7.8** (3.06)
Conditional	2.41 (8.87)	-14.1 (10.1)	2.53 (3)	-13.6** (6.08)	5.53 (6.44)	-6.06** (2.84)
Mean	58.088	88.556	23.224	51.004	79.601	23.242
Identified set	[2.415, 6.143]†	[-17.227, -14.104]†	[1.034, 2.526]†	[-13.650, -12.787]†	[4.712, 5.527]†	[-6.061, -5.241]†
δ for 0 coefficient	-0.697	-5.768	1.604	9.540	5.597	4.694
R^2 (unc., con.)	(0.035, 0.400)	(0.035, 0.386)	(0.035, 0.096)	(0.020, 0.283)	(0.020, 0.364)	(0.020, 0.074)
<i>Outcome: Leisure (in minutes)</i>						
Unconditional	2.06 (6.27)	4.48 (11.6)	-13.1* (7.92)	-6.77 (6.16)	-1.32 (10.1)	4.64 (5.97)
Conditional	5.11 (5.22)	4.06 (9.69)	-11.5 (7.73)	-6.51 (5.46)	-6.58 (8)	4.21 (5.87)
Mean	20.756	76.831	108.088	27.482	84.705	106.950
Identified set	[5.112, 6.324]†	[3.909, 4.065]†	[-11.533, -10.825]†	[-6.508, -6.411]†	[-8.451, -6.577]†	[4.005, 4.205]†
δ for 0 coefficient	-5.450	12.126	7.753	19.510	-4.115	10.963
R^2 (unc., con.)	(0.019, 0.212)	(0.019, 0.452)	(0.019, 0.099)	(0.026, 0.195)	(0.026, 0.429)	(0.026, 0.110)
<i>Outcome: Sleep (in minutes)</i>						
Unconditional	-8.81** (4.43)	-.704 (5.39)	3.87 (10.3)	4.18 (2.73)	1.05 (3.58)	-5.19 (7.91)
Conditional	-5.94 (3.86)	-1.11 (5.41)	2.64 (9.54)	3.81 (2.88)	-2.48 (3.08)	-11.3 (7.69)
Mean	9.170	15.100	460.740	8.612	13.419	463.357
Identified set	[-5.943, -4.672]†	[-1.272, -1.112]†	[1.921, 2.644]†	[3.661, 3.810]†	[-3.706, -2.482]†	[-14.057, -11.293]†
δ for 0 coefficient	3.621	-11.203	3.078	11.843	-2.210	-4.592
R^2 (unc., con.)	(0.066, 0.112)	(0.066, 0.253)	(0.066, 0.178)	(0.032, 0.077)	(0.032, 0.270)	(0.032, 0.130)
Observations	351	351	351	507	507	507

Notes: Center hours are from 8am-4pm on weekdays, evening and weekend consists of 4pm-8pm on weekdays and entire weekend days (8am-8pm). Nights are from 8pm-8am. Table shows coefficients from OLS regressions of the outcome variables on an indicator variable for center-based care usage. Figure 5.4 shows the conditional coefficients and the coefficient under the assumption of equally large selection on observables as on unobservables ($\delta = 1$). See Table 5.2 for other table notes and section 5.4 for details on the empirical specification. Source: German Time-Use Survey (2012/13)

Table 5.A.10: The effect of full-day care on parenting activities (Oster bounds)

	Households with lower maternal education			Households with higher maternal education		
	All parents (1)	Mothers (2)	Fathers (3)	All parents (4)	Mothers (5)	Fathers (6)
<i>Outcome: Reading books or telling stories (daily)</i>						
Unconditional	-0.049*** (0.019)	-0.047** (0.022)	-0.042 (0.031)	0.003 (0.015)	-0.003 (0.016)	0.001 (0.025)
Conditional	-.0529*** (.0185)	-.0457** (.0232)	-.0735** (.0314)	-.000586 (.0146)	-.00389 (.0158)	-.0124 (.0263)
Mean	0.615	0.718	0.455	0.773	0.879	0.645
Identified set	[-0.054, -0.053]†	[-0.046, -0.045]†	[-0.085, -0.074]†	[-0.002, -0.001]†	[0.004, 0.006]†	[-0.017, -0.012]†
δ for 0 coefficient	73.042	19.413	-8.148	-0.534	-1.893	-3.065
R^2 (unc., con.)	(0.000, 0.155)	(0.000, 0.085)	(0.000, 0.122)	(0.000, 0.130)	(0.000, 0.039)	(0.000, 0.100)
<i>Outcome: Musical activities (daily)</i>						
Unconditional	-0.006 (0.017)	0.018 (0.023)	-0.034 (0.021)	-0.012 (0.017)	-0.029 (0.024)	-0.002 (0.021)
Conditional	-.00906 (.0173)	.0116 (.0244)	-.038* (.0216)	-.0176 (.0167)	-.0155 (.0246)	-.0248 (.0213)
Mean	0.254	0.326	0.136	0.320	0.436	0.180
Identified set	[-0.010, -0.009]†	[0.009, 0.012]†	[-0.039, -0.038]†	[-0.019, -0.018]†	[-0.015, -0.010]†	[-0.033, -0.025]†
δ for 0 coefficient	-10.709	4.424	61.506	-15.015	2.816	-3.199
R^2 (unc., con.)	(0.002, 0.064)	(0.002, 0.024)	(0.002, 0.043)	(0.002, 0.109)	(0.002, 0.065)	(0.002, 0.044)
<i>Outcome: Painting, building or drawing (daily)</i>						
Unconditional	-0.029* (0.017)	-0.040* (0.024)	-0.002 (0.023)	-0.043*** (0.016)	-0.107*** (0.023)	0.031 (0.021)
Conditional	-.0202 (.0171)	-.0329 (.024)	.000965 (.0232)	-.0438*** (.0162)	-.083*** (.0231)	.0037 (.0222)
Mean	0.287	0.367	0.161	0.274	0.348	0.184
Identified set	[-0.020, -0.017]†	[-0.033, -0.030]†	[0.001, 0.002]†	[-0.044, -0.044]†	[-0.083, -0.073]†	[-0.006, 0.004]
δ for 0 coefficient	5.542	8.429	-0.925	31.544	5.675	0.383
R^2 (unc., con.)	(0.003, 0.119)	(0.003, 0.089)	(0.003, 0.063)	(0.003, 0.104)	(0.003, 0.108)	(0.003, 0.063)
<i>Outcome: Playing games together (daily)</i>						
Unconditional	-0.091*** (0.019)	-0.078*** (0.024)	-0.104*** (0.029)	-0.057*** (0.018)	-0.122*** (0.023)	0.014 (0.026)
Conditional	-.1*** (.0188)	-.0911*** (.0245)	-.12*** (.0299)	-.0578*** (.0177)	-.125*** (.0234)	.0189 (.0265)
Mean	0.479	0.553	0.364	0.516	0.594	0.422
Identified set	[-0.104, -0.100]†	[-0.096, -0.091]†	[-0.126, -0.120]†	[-0.058, -0.058]†	[-0.126, -0.125]†	[0.019, 0.021]†
δ for 0 coefficient	198.916	-60.429	-428.016	29.760	21.874	-18.574
R^2 (unc., con.)	(0.002, 0.127)	(0.002, 0.107)	(0.002, 0.098)	(0.002, 0.129)	(0.002, 0.114)	(0.002, 0.122)
<i>Outcome: Outdoor activities (daily)</i>						
Unconditional	-0.060*** (0.019)	-0.082*** (0.024)	-0.012 (0.028)	-0.022 (0.018)	-0.072*** (0.023)	0.026 (0.025)
Conditional	-.0727*** (.0187)	-.1*** (.0244)	-.027 (.0293)	-.0562*** (.0173)	-.0903*** (.0236)	-.0138 (.0255)
Mean	0.487	0.606	0.301	0.484	0.623	0.316
Identified set	[-0.077, -0.073]†	[-0.107, -0.100]†	[-0.033, -0.027]†	[-0.068, -0.056]†	[-0.097, -0.090]†	[-0.028, -0.014]†
δ for 0 coefficient	-35.182	-27.670	-6.188	-5.491	-21.183	-1.027
R^2 (unc., con.)	(0.002, 0.149)	(0.002, 0.100)	(0.002, 0.046)	(0.002, 0.170)	(0.002, 0.100)	(0.002, 0.090)
<i>Outcome: Gymnastics, sports (daily)</i>						
Unconditional	0.007 (0.019)	0.014 (0.024)	-0.007 (0.030)	0.005 (0.017)	0.010 (0.023)	0.001 (0.026)
Conditional	-.0157 (.0195)	-.0034 (.025)	-.0338 (.0316)	-.00991 (.0178)	.00543 (.0239)	-.0357 (.0267)
Mean	0.406	0.399	0.417	0.386	0.349	0.431
Identified set	[-0.023, -0.016]†	[-0.010, -0.003]†	[-0.043, -0.034]†	[-0.015, -0.010]†	[0.004, 0.005]†	[-0.049, -0.036]†
δ for 0 coefficient	-2.195	-0.573	-4.041	-2.100	3.015	-3.110
R^2 (unc., con.)	(0.000, 0.051)	(0.000, 0.051)	(0.000, 0.068)	(0.000, 0.076)	(0.000, 0.046)	(0.000, 0.130)
<i>Outcome: Watching television or videos (daily)</i>						
Unconditional	-0.019 (0.019)	-0.030 (0.024)	0.002 (0.031)	0.024 (0.018)	0.011 (0.024)	0.038 (0.026)
Conditional	-.0302 (.0197)	-.0442* (.0252)	-.0112 (.032)	.00941 (.0186)	-.00116 (.0254)	.0243 (.0274)
Mean	0.518	0.564	0.446	0.435	0.459	0.406
Identified set	[-0.034, -0.030]†	[-0.049, -0.044]†	[-0.016, -0.011]†	[0.004, 0.009]†	[-0.006, -0.001]†	[0.019, 0.024]†
δ for 0 coefficient	-10.951	-11.421	-2.747	1.754	-0.280	3.984
R^2 (unc., con.)	(0.002, 0.053)	(0.002, 0.044)	(0.002, 0.057)	(0.002, 0.040)	(0.002, 0.037)	(0.002, 0.053)
Observations	2852	1757	1095	3114	1712	1402

Notes: Table shows coefficients from OLS regression of the outcome variables (binary indicator that equals one if the activity is being performed daily) on an indicator variable for full-day care. Unconditional coefficients stem from a regression which only includes the full-day indicator and dummies for child age. The conditional coefficients are from regressions that additionally include wave dummies and the set of controls described in the table notes of Table 5.3. See Table 5.2 and section 5.4 for other notes on the Oster-method. Source: pairfam, 2013-2019.

Table 5.A.11: The effect of full-day care on non-parenting activities and other parent- and child-related outcomes; pairfam - Oster bounds

	Households with lower maternal education			Households with higher maternal education		
	All parents (1)	Mothers (2)	Fathers (3)	All parents (4)	Mothers (5)	Fathers (6)
<i>Outcome: Working (at least 10 h/w)</i>						
Unconditional	0.113*** (0.017)	0.161*** (0.023)	0.022 (0.016)	0.105*** (0.014)	0.201*** (0.021)	-0.002 (0.012)
Conditional	.116*** (.0161)	.173*** (.0238)	.0275 (.0169)	.105*** (.0139)	.185*** (.0225)	.000318 (.0134)
Mean	0.684	0.571	0.866	0.801	0.684	0.943
Identified set	[0.116, 0.118]†	[0.173, 0.179]†	[0.027, 0.029]†	[0.105, 0.105]†	[0.174, 0.185]†	[0.000, 0.001]†
δ for 0 coefficient	30.807	32.380	-41.228	18.164	4.449	-0.380
R ² (unc., con.)	(0.000, 0.199)	(0.000, 0.096)	(0.000, 0.111)	(0.000, 0.146)	(0.000, 0.099)	(0.000, 0.036)
<i>Outcome: Working hours (per week)</i>						
Unconditional	4.509*** (0.717)	7.053*** (0.747)	-0.617 (0.888)	4.919*** (0.634)	9.824*** (0.719)	-0.235 (0.740)
Conditional	3.83*** (.581)	6.16*** (.749)	.466 (.915)	4.38*** (.547)	8.24*** (.751)	-.251 (.787)
Mean	24.297	15.751	38.032	30.077	20.721	41.369
Identified set	[3.584, 3.827]†	[5.605, 6.160]†	[0.466, 0.862]†	[4.184, 4.382]†	[6.735, 8.243]†	[-0.256, -0.251]†
δ for 0 coefficient	9.790	4.201	-1.272	11.093	2.528	51.367
R ² (unc., con.)	(0.000, 0.436)	(0.000, 0.117)	(0.000, 0.094)	(0.000, 0.358)	(0.000, 0.152)	(0.000, 0.039)
<i>Outcome: Personal monthly net income</i>						
Unconditional	107.146** (44.057)	265.750*** (38.633)	-221.252*** (71.421)	112.236* (64.929)	462.932*** (49.950)	-271.337** (118.409)
Conditional	165*** (33.2)	261*** (37.6)	27.5 (58.2)	229*** (60)	420*** (51.1)	50.7 (126)
Mean	1124.115	611.010	1949.043	1807.412	1099.077	2677.208
Identified set	[164.648, 184.670]†	[258.398, 260.818]†	[27.489, 115.481]†	[228.833, 269.982]†	[396.752, 420.413]†	[50.669, 171.711]†
δ for 0 coefficient	-11.717	9.829	-0.325	-6.926	5.566	-0.433
R ² (unc., con.)	(0.013, 0.457)	(0.013, 0.106)	(0.013, 0.330)	(0.013, 0.265)	(0.013, 0.134)	(0.013, 0.136)
<i>Outcome: Too little time with child (0/1)</i>						
Unconditional	0.076*** (0.023)	0.113*** (0.028)	0.007 (0.038)	0.092*** (0.021)	0.185*** (0.026)	-0.016 (0.032)
Conditional	.0661*** (.0239)	.0945*** (.0292)	.0255 (.0414)	.0824*** (.0217)	.187*** (.0281)	-.0495 (.0335)
Mean	0.360	0.276	0.498	0.397	0.303	0.510
Identified set	[0.062, 0.066]†	[0.085, 0.095]†	[0.025, 0.032]†	[0.079, 0.082]†	[0.187, 0.188]†	[-0.061, -0.049]†
δ for 0 coefficient	8.640	4.314	-4.600	9.362	5.789	-4.553
R ² (unc., con.)	(0.001, 0.083)	(0.001, 0.049)	(0.001, 0.030)	(0.001, 0.079)	(0.001, 0.077)	(0.001, 0.055)
<i>Outcome: Feeling stressed (1-5)</i>						
Unconditional	-0.022 (0.045)	0.061 (0.056)	-0.142* (0.074)	0.152*** (0.039)	0.263*** (0.050)	0.007 (0.060)
Conditional	.00643 (.0472)	.052 (.0595)	-.0871 (.0796)	.145*** (.0407)	.265*** (.0522)	-.00481 (.0645)
Mean	3.229	3.316	3.087	3.296	3.367	3.210
Identified set	[0.006, 0.016]†	[0.053, 0.055]†	[-0.087, -0.066]†	[0.142, 0.145]†	[0.265, 0.267]†	[-0.009, -0.005]†
δ for 0 coefficient	-0.685	12.653	3.379	11.033	7.319	-1.215
R ² (unc., con.)	(0.000, 0.031)	(0.000, 0.019)	(0.000, 0.037)	(0.000, 0.027)	(0.000, 0.034)	(0.000, 0.025)
<i>Outcome: Hours of sleep (parent)</i>						
Unconditional	-0.034 (0.044)	-0.045 (0.058)	-0.014 (0.067)	-0.013 (0.034)	-0.050 (0.047)	0.025 (0.048)
Conditional	-.00699 (.0469)	-.0000499 (.0605)	-.0348 (.0751)	.00322 (.0355)	-.0454 (.0498)	.0495 (.0495)
Mean	6.771	6.807	6.713	6.840	6.862	6.814
Identified set	[-0.007, 0.003]†	[-0.000, 0.016]†	[-0.043, -0.035]†	[0.003, 0.009]†	[-0.045, -0.044]†	[0.050, 0.059]†
δ for 0 coefficient	0.741	0.003	-5.369	-0.611	14.202	-6.855
R ² (unc., con.)	(0.003, 0.036)	(0.003, 0.046)	(0.003, 0.040)	(0.003, 0.034)	(0.003, 0.055)	(0.003, 0.040)
<i>Outcome: Hours of sleep (child)</i>						
Unconditional	-0.255*** (0.048)	-0.274*** (0.062)	-0.224*** (0.074)	-0.190*** (0.038)	-0.208*** (0.049)	-0.170*** (0.059)
Conditional	-.181*** (.0506)	-.2*** (.0665)	-.138* (.0782)	-.131*** (.041)	-.154*** (.0539)	-.112* (.064)
Mean	10.323	10.321	10.325	10.461	10.526	10.384
Identified set	[-0.181, -0.147]†	[-0.200, -0.167]†	[-0.138, -0.100]†	[-0.131, -0.105]†	[-0.154, -0.129]†	[-0.112, -0.088]†
δ for 0 coefficient	3.813	4.118	2.981	3.724	3.943	3.657
R ² (unc., con.)	(0.008, 0.045)	(0.008, 0.053)	(0.008, 0.052)	(0.008, 0.041)	(0.008, 0.039)	(0.008, 0.052)

continues on next page

continued from previous page

<i>Outcome: Child is happy and content (1-5)</i>						
Unconditional	-0.064** (0.026)	-0.090*** (0.034)	-0.025 (0.041)	0.031 (0.022)	0.002 (0.029)	0.065** (0.033)
Conditional	-.0734*** (.0278)	-.0963*** (.0366)	-.0337 (.0427)	.0396* (.023)	.017 (.0312)	.0631* (.034)
Mean	4.536	4.534	4.539	4.552	4.581	4.518
Identified set	[-0.077, -0.073]†	[-0.099, -0.096]†	[-0.037, -0.034]†	[0.040, 0.043]†	[0.017, 0.022]†	[0.062, 0.063]†
δ for 0 coefficient	-74.897	31.341	-17.398	-23.505	-3.782	15.605
R^2 (unc., con.)	(0.003, .0127)	(0.003, .0213)	(0.003, .0213)	(0.003, .0194)	(0.003, .0234)	(0.003, .045)
<i>Outcome: Child is irritable and cries often (1-5)</i>						
Unconditional	-0.053 (0.043)	-0.011 (0.057)	-0.112* (0.063)	-0.052 (0.037)	-0.006 (0.050)	-0.104* (0.054)
Conditional	-.0321 (.0444)	-.0373 (.0592)	-.0261 (.0668)	-.0335 (.0386)	-.00797 (.0524)	-.0733 (.0576)
Mean	2.295	2.326	2.247	2.226	2.147	2.318
Identified set	[-0.032, -0.025]†	[-0.047, -0.037]†	[-0.026, 0.005]	[-0.033, -0.027]†	[-0.009, -0.008]†	[-0.073, -0.061]†
δ for 0 coefficient	3.820	-4.579	0.838	4.346	-15.174	4.700
R^2 (unc., con.)	(0.003, .0213)	(0.003, .018)	(0.003, .048)	(0.003, .0296)	(0.003, .0276)	(0.003, .0346)
Observations	2859	1763	1096	3135	1725	1410

Notes: Table shows coefficients from OLS regressions of the outcome variables on an indicator variable for full-day care. Unconditional coefficients stem from a regression which only includes the full-day indicator and dummies for child age. The conditional coefficients are from which regressions that additionally include wave dummies and the set of controls described in the table notes of Table 5.3. See Table 5.2 and section 5.4 for further notes on the Oster-method. Source: pairfam, 2013-2019.

Table 5.A.12: The effect of full-day care on parenting activities - alternative full-day assignment

	Households with lower maternal education			Households with higher maternal education		
	All parents (1)	Mothers (2)	Fathers (3)	All parents (4)	Mothers (5)	Fathers (6)
<i>Panel A: Parenting activities</i>						
Reading books or telling stories (daily)	-0.021 (0.023)	-0.035 (0.028)	-0.018 (0.039)	-0.013 (0.016)	-0.001 (0.016)	-0.026 (0.030)
Singing or playing instruments (daily)	-0.004 (0.022)	0.000 (0.032)	0.002 (0.030)	-0.018 (0.020)	-0.025 (0.029)	-0.023 (0.026)
Painting, building or drawing (daily)	-0.019 (0.023)	-0.014 (0.033)	-0.020 (0.033)	-0.023 (0.020)	-0.058** (0.028)	0.010 (0.027)
Playing games together (daily)	-0.082*** (0.025)	-0.069** (0.032)	-0.108*** (0.041)	-0.021 (0.021)	-0.054** (0.027)	0.027 (0.032)
Outdoor activities (daily)	-0.073*** (0.024)	-0.094*** (0.031)	-0.027 (0.040)	-0.074*** (0.020)	-0.150*** (0.027)	0.019 (0.031)
Gymnastics, sports (daily)	-0.020 (0.025)	0.006 (0.032)	-0.059 (0.042)	-0.042** (0.021)	-0.050* (0.028)	-0.033 (0.032)
Watching television or videos (daily)	0.006 (0.026)	-0.016 (0.033)	0.042 (0.042)	0.038* (0.022)	-0.002 (0.030)	0.090*** (0.032)
<i>Panel B: Non-parenting activities and other outcomes</i>						
Working (at least 10 h/w)	0.142*** (0.021)	0.204*** (0.030)	0.044* (0.023)	0.136*** (0.016)	0.270*** (0.025)	-0.026* (0.015)
Working hours (per week)	6.296*** (0.747)	9.084*** (0.962)	1.993* (1.210)	5.263*** (0.620)	11.496*** (0.814)	-2.215** (0.890)
Personal monthly net income	256.669*** (42.882)	409.269*** (50.097)	23.678 (75.848)	222.188*** (66.036)	559.758*** (57.916)	-168.564 (134.647)
Too little time with child (0/1)	0.076*** (0.027)	0.081** (0.033)	0.065 (0.046)	0.120*** (0.023)	0.221*** (0.028)	-0.006 (0.035)
Feeling stressed (1-5)	0.092 (0.061)	0.058 (0.077)	0.135 (0.101)	0.170*** (0.047)	0.300*** (0.061)	0.012 (0.075)
Hours of sleep (parent)	-0.134** (0.059)	-0.106 (0.076)	-0.186* (0.097)	-0.048 (0.042)	-0.111* (0.058)	0.037 (0.060)
Hours of sleep (child)	-0.290*** (0.064)	-0.252*** (0.084)	-0.312*** (0.101)	-0.240*** (0.045)	-0.185*** (0.056)	-0.317*** (0.074)
Child is happy and content (1-5)	-0.064* (0.036)	-0.081* (0.045)	-0.067 (0.060)	-0.018 (0.026)	-0.048 (0.034)	0.019 (0.040)
Child is irritable and cries often (1-5)	0.012 (0.056)	-0.004 (0.073)	0.056 (0.091)	0.092** (0.044)	0.161*** (0.060)	0.007 (0.066)
Observations	1972	1209	763	2338	1295	1043

Notes: Table shows coefficients from OLS regressions of the outcome variables (binary indicator that equals one if the activity is being performed daily) on an indicator variable for full-day care. Additional controls; dummies for child age, number of children in family, parental sex (if applicable), age of parent, indicator for migration status, single parent indicator, education dummies (if applicable). Full-day care indicates whether the child attends center-based care 30+ vs. 15-30 hours per week. See Table 5.3 for other table notes. Source: pairfam, 2013-2019.

CHAPTER 6

When equality doesn't make it indoors: Gender differences in market and non-market work in the GDR and reunified Germany¹

6.1 Introduction

Gender (in)equality in market and non-market work is continuing to attract attention in academia and in the wider public. Rising female labour force participation has been one of the central and most-studied phenomena in labour economics (Altonji and Blank, 1999; Bertrand, 2020; Goldin, 2014; Olivetti and Petrongolo, 2016). While substantial progress was made in the last decades, important differences in labour supply remain. At the OECD-level, about 80% of part-time employment is by females, and gender-gaps in income remain: women working full-time earn about 13% lower wages than men.² At least since Becker's *A Treatise on the Family* (Becker, 1981), such differences in market work are seen as inherently linked to non-market work. In standard non-cooperative bargaining models, two individuals maximise utility taking into account their own and their partner's labour market returns (Lundberg and Pollak, 1996; Vermeulen, 2002). A typical implication of these models is that labour market inequality directly affects inequality also in non-market work. In short, if women cannot earn (the same amount of) independent income, they will also have to do more of the unpleasant household tasks, and have less power to make household-level decisions generally.³

¹This chapter is joint work with Jan Berkes (DIW Berlin and Freie Universität Berlin) and Felix Weinhardt (Viadrina European University, DIW Berlin, IZA, CEP and CESifo). We are grateful to Ludovica Gambaro and C. Katharina Spiess, as well as seminar participants at DIW Berlin and ifo Munich for helpful comments.

²Data from <https://data.oecd.org/earnwage/gender-wage-gap.htm>, accessed July 21, 2021.

³Individuals may experience different disutility from paid work and housework (e.g., due to the stress or reward associated with the work). Combined with positive, increasing returns to specialisation, corner solutions as a result of the bargaining process may exist with each partner being solely responsible for

This paper takes this hypothesis to a test: does the amount of individual housework contribution depend on their spouses' labour supply? Our answer, based on data from different regimes and over time, is; *no*. Individual housework can almost entirely be explained by one's own time spent in paid work.

This is surprising, also because descriptive evidence suggests otherwise. Using multinational time-use data, Figure 6.1 shows the relation between the within-couple female share in market work and housework across 73,214 households, for 16 countries, covering a time-span from 1974 to 2014.⁴ Female time-shares of market and household-work are pronounced inversely related, showing almost a negative one-to-one relationship. For example, we see the more egalitarian Scandinavian countries towards the lower right end, where women spend a higher share of their time in the labour market and equally a lower share doing household tasks.⁵ Similarly, this relationship also holds for countries over time where the distribution in both domains becomes more equal over time (e.g., West Germany). This descriptive evidence could be taken as support of the (theoretical) view that gender equality in non-market work is closely linked to a levelling of the playing field for market work.

Providing a causal interpretation of the relation shown in Figure 6.1 is non-trivial. Underlying is a complicated process of (potentially joint) decision making, with additional constraints or preference parameters imposed by societal norms or institutional settings. These might interact with third (endogenous) factors, such as the presence of children. Moreover, the partner choice itself is potentially endogenous to the later-observed household-level realisation of time-allocation (after accounting for selection into employment).

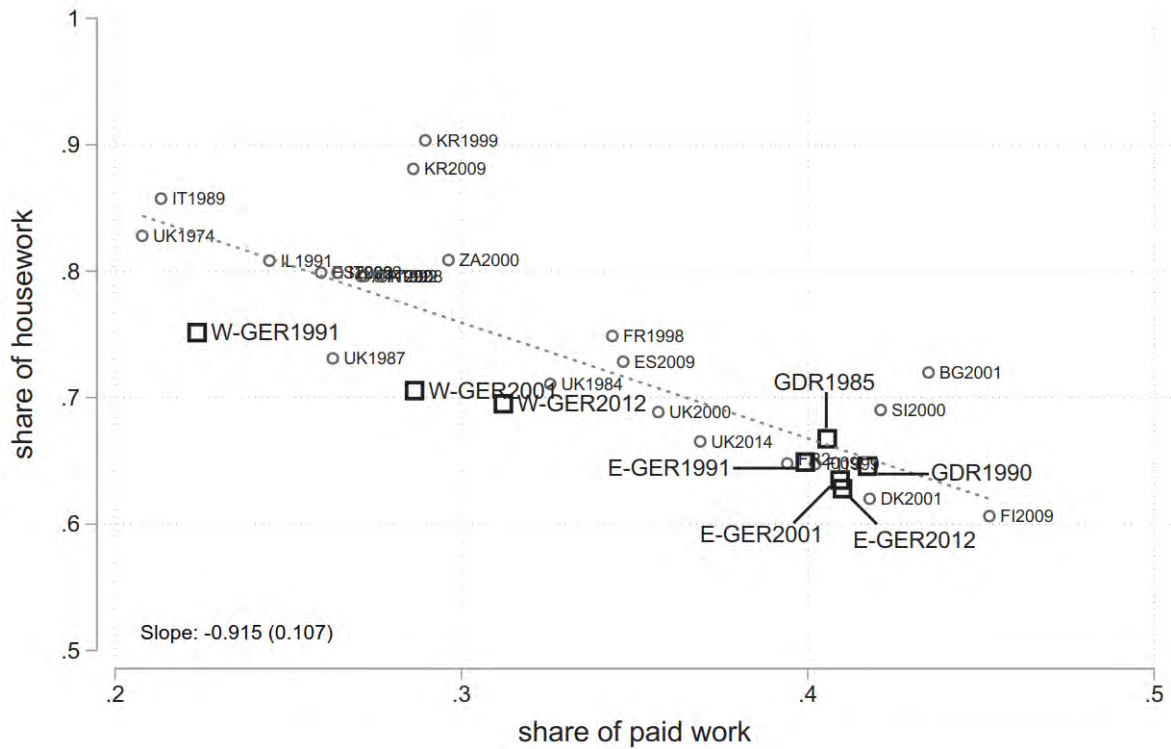
To circumvent these challenges, we study household-level time allocation before and after German reunification, making use of five time-use surveys from 1985 to 2013. These include newly sourced time-use data from the GDR, which to the best of our knowledge have not been analysed by economists before. The combination of data availability and the institutional setting of the GDR and West Germany is of particular interest. Due to different institutions, women in the GDR spent much more time in the labour market than their Western counterparts. These institutional differences shifted female labour supply out in the GDR in ways arguably not related to the intra-household decision process regarding time spent doing housework. This is because in the GDR women had limited choices regarding labour supply decisions, and were

market work and housework (*separate spheres equilibrium*, see Lundberg and Pollak, 1993). However, this does not explain the common distribution that mostly women do a large share of housework.

⁴See Appendix Table 6.A.1 for a list of countries, survey years, and observations from the Multinational Time Use Study (Gershuny et al., 2020).

⁵Household tasks exclude child care and care for elderly and are defined in section 6.3.

Figure 6.1: Female share of housework and paid work



Notes: Figure plots the within-couple female share of housework and paid work in a cross-country comparison. Data from GDR time budget study (GDR), German time use data of East Germany (E-GER) and West Germany (W-GER) and the Multinational Time Use Study (other countries). Sample is restricted to different-sex couples aged 18-65 and surveys with at least 400 couple-level observations.

downward-constrained as full participation was mandated and wages were largely fixed by the government (Krueger and Pischke, 1995). On the other hand, women in West Germany were able to make choices themselves. However, it is well-documented that many women in West Germany were upward-constrained in their choices due to norms and limitations related to the child care sector (Rosenfeld et al., 2004). We describe these institutional differences in greater detail below. Notably, we can not only study resulting (partly quasi-exogenous) differences in household-level labour supply between the GDR and West Germany, but also track these in reunified Germany into present times. Since the two German states represent data points on the opposite end of Figure 6.1, we believe our results also allow drawing general implications beyond the German context.

Based on descriptive and regression analyses, this paper proceeds in three steps.

We first provide descriptive evidence on the time allocated to market and non-market work in the GDR, as well as East and West Germany in 1991/92, i.e. shortly after

reunification. This analysis confirms that it is true that women in the GDR worked more hours in the labour market compared to women in the West, and spent fewer hours on household tasks. Similarly, the distribution of market work and housework is more equal in East compared to West Germany.

In the second step, we decompose the relation between female shares in market work and housework by household type. To do this, we classify households into three types, i.e. (main) male breadwinner, dual earner, and (main) female breadwinner, based on the time spent on market work. While the male breadwinner is the norm in the West, the dual earner is the norm in the GDR and East Germany. Accounting for these underlying types, the female shares in housework look remarkably similar across countries/regions, and in fact show a negative relation comparable to the cross-country evidence presented in Figure 6.1. One explanation for these patterns is that there are no male reactions to different levels in female market work regarding their household work, and vice versa, implying that the greater equality documented in the first step of the analysis is not because of behavioural differences but a mechanical effect of individual constraints.

In the third step, we use a decomposition to test this hypothesis directly. The key finding from this analysis is that once individual labour supply is accounted for, the gender housework gap between East and West Germany is reduced by 89%. We then investigate heterogeneity and find that this result is not driven by the presence of children. In fact, households without children provide the cleaner analysis since complementarities between child-rearing and housework can be excluded, and this is where remaining gaps between the East and West are the smallest. Our headline finding is therefore that once selection into employment is accounted for, the distribution of housework looks remarkably similar across the two German states.

Taken together, these findings have important implications. First, institutional changes that increase women's employment (and reduce income differences) are unlikely to affect the relation of time-allocation between market work and housework. Such policies should be seen as, at best, reducing gaps in the labour market but not be expected to shift women's time spent on housework to a fully gender-equal level or cause spouses to relieve their wives from their domestic burden. At least the extreme labour market policies in the GDR for a duration of four decades did not make partners' time allocation to housework more responsive to their spouses' labour supply, and thus lead to gender inequality in the domestic domain. Second, our results put into perspective the narrative about higher gender equality in the GDR. While it is beyond doubt that women in the GDR had a higher labour market attachment, compared to the West, we do not find evidence for behavioural differences in the allocation of

housework. Third, our findings shed light onto the nature of gender norms, as we show that the relation between market and household work is primarily driven by individual constraints.

Our study relates to several strands of literature. First of all, we contribute to the literature examining how East and West Germans differ in gender-related attitudes due to exposure to different policies during the division. Campa and Serafinelli (2019) use SOEP survey data to provide evidence that East German women place more importance on career success (see also Beblo and Görge, 2018). Using time-use data, we also see that an equal distribution of paid work *within* couples continues to be more common in East Germany. Lippmann et al. (2020) show that West German wives—but not East German ones—are more likely to increase their housework contribution and to withdraw from the labour market after outearning their husbands in order to conform with traditional gender roles. Only in West German couples is the risk of divorce increased when the wife is earning more. Zoch (2021) analyses East-West differences in attitudes towards maternal employment and housework and finds pronounced disparities in attitudes, but those have become smaller for younger cohorts.⁶ Our findings imply that such differences—including those on housework norms—are likely to be primarily driven by higher (state-imposed) female labour force participation in the GDR and that norms on housework were not affected independently of this.

Moreover, our empirical results speak to the literature on the relation between labour markets and gender norms (Grunow et al., 2018). Our results demonstrate that if preferences and social norms about housework evolve isolated from other domains of gender equality such as market work (Hakim, 2000), policies can have unintended, even detrimental, consequences for women. One explanation for the weak link between equality in the labour market and in domestic spheres is that female labour force participation is publicly visible which allows for local learning to take place gradually over time (Fernández, 2013; Fogli and Veldkamp, 2011). In contrast, housework is rarely observed beyond families which might slow down learning processes.

Last but not least, we complement recent findings that males do not adjust their own labour supply when their wives start working (Knowles, 2013). We show that housework also does not adjust depending on the partner's labour supply. Both of these findings suggest that partner-reactions in the non-cooperative household bargaining framework are small.

⁶See also Bauernschuster and Rainer (2012) who document large differences in attitudes towards maternal employment and the role of wives in the family immediately following reunification.

6.2 Institutional setting

After the end of the atrocities of World War II, Germany was divided into four occupation zones by the victors in 1945. In 1949, the GDR was formally established in the Soviet occupation zone and the FRG consisting of the three western zones. The GDR was a socialist, one-party state under strong influence of the Soviet Union. In contrast, in the FRG a market-based democracy was established.

During 41 years of formal separation the two German states diverged in many regards, including female labour force participation (Trappe, 1996). The GDR enforced high participation rates through several policies. E.g., child care was strongly expanded⁷ and not working was considered to be anti-social behaviour (Beblo and Görge, 2018). As a result, female labour force participation increased strongly and in 1989 with a rate of 89% it was among the highest in the world. Additionally, most women worked full-time and differences by marital status and children was small (Rosenfeld et al., 2004).

Gender policies in the FRG were conservative in comparison. Limited child care availability and afternoon care as well as joint taxation for married couples favoured (main) male breadwinner households (Boelmann et al., 2021). Female labour force participation was a third lower than men's and part-time work was prevalent, especially so for young mothers.⁸ A popular children's song in the GDR was "Wenn Mutti früh zur Arbeit geht" ("*When mommy goes to work in the morning*"), whereas in the FRG wives by law only had "[...] the right to be employed as far as this is compatible with her marriage and family duties" until 1977 (Lippmann et al., 2020). Despite these differences, gender earning gaps were about similar with 25% for full-time workers (Krueger and Pischke, 1995), but as outlined above, selection into employment was much lower for women in the GDR. Similarly, wage gaps differed relatively little with 15% in the GDR and 18% in the FRG (Sørensen and Trappe, 1995).

In summary, gaps in hourly wages were comparable, but gaps in labour supply large. We use this setting to study how resulting differences in earnings shift time-allocation to unpleasant household-tasks across these regimes. Notably, mirroring the institutionally-induced differences, we study the relation between labour supply and time spent doing household tasks.⁹

⁷In 1989, 98% of children aged 3-6 attended child care facilities and more than 80% of children below 3 (Schmitz and Weinhardt, 2019). In contrast, in the FRG child care for under threes was basically non-existent and for older children almost all spots were part-time only.

⁸In the GDR it was mostly older women working reduced hours.

⁹Pollak (2005) argues that what matters for household bargaining are not realised earnings but earnings potentials (i.e. wages). But this assumes that the hours worked are a choice variable, which is precisely not the case in the institutional settings studied here.

The two German states were reunified in October 1990 following the fall of the Berlin Wall one year before. East Germany fully adapted the policies of the FRG, with arguably the most notable difference remaining being the higher provision of child care spots in East Germany.

6.3 Data

GDR For our analysis we obtained access to the 1985 and 1990 waves of the GDR time budget study (*Zeitbudgeterhebung*) at the German Federal Archives. The study was conducted by the statistical office of the GDR to obtain data for the planning of demand for goods and services, to demonstrate the effectiveness of economic and social policies on the use of time outside of work and to design new reforms that foster efficient time-use (Fiebiger, 1991).

The 1985 wave documents time-use in the GDR years before the fall of the wall in 1989. The 1990 wave was collected before Germany was officially reunified into a monetary, economic and social union. Data collection of the GDR time budget study only took place among worker, employee, and retiree households.¹⁰ Each household was supposed to fill out the survey on a pre-determined day of the week. Main tasks were documented for 24 hours, starting at midnight.

Reunified Germany We use three waves of the German Time-Use Survey from after reunification. The study contains around 5,000 households in each survey wave taken in 1991/92, 2001/02 and 2012/13 (Maier, 2014). Each adult household member records their activities in ten minutes slots over three survey days (five minutes over two days in 1991/92). The activities are categorised in detailed three-digit activities. Besides the diary data for the survey days, the data also contain other household and individual characteristics. Throughout the analysis we treat East and West Germany separately and we mostly focus on the 1991/92 study conducted briefly after reunification.

Harmonisation of data sets and sample selection For our analysis we harmonise the time use studies to make them directly comparable. First, we define consistent categories of activities. The broad categories we are interested in are:¹¹

- paid work
- housework

¹⁰Priller (1993) confirms that the data is representative for worker and employee households by districts but that one-person households and young male respondents are slightly underrepresented. Since we study couple households this is of less concern for our study.

¹¹Larger other categories are leisure, sleep, eating, personal hygiene, and care for adults. Care for adults comprises less than 8% of the category *care for others* and is only conducted by 4% of individuals. Accordingly, child care is quantitatively much more relevant.

- child care

A detailed list of activities contributing to the broader categories are presented in Appendix Table 6.A.2. As our analysis is conducted on a household level, we impose some sample restrictions; we look at (married) couples and due to the focus on gender differences restrict this to different-sex couples. As we look at the interplay of time invested in paid work and housework, we further restrict the couples to be of working age, i.e. 18-65. We focus on weekdays which mostly are regular working days with a positive number of working hours. The upper panel of Table 6.1 shows characteristics of our analysis sample.

Table 6.1: Summary statistics of time-use data

	GDR (85 and 90)		East Germany (91/92)		West Germany (91/92)	
	Women	Men	Women	Men	Women	Men
<i>Characteristics</i>						
Age	38.94	41.22	40.88	43.28	42.49	45.57
Employed	0.92	0.98	0.68	0.80	0.62	0.91
High vocational degree	0.34	0.32
Upper secondary school	.	.	0.25	0.30	0.19	0.30
Children under 10 years in household	0.44	0.44	0.40	0.40	0.39	0.39
Weekday (Mo-Fr)	0.72	0.72	0.78	0.78	0.75	0.75
<i>Time use in minutes (weekday)</i>						
Paid work (total)	429.27	556.55	325.87	462.42	157.64	467.21
Housework	223.78	121.89	250.55	148.86	325.55	116.53
Child care	41.59	11.80	47.87	18.23	59.10	16.91
Observations	3237	3237	2154	2154	6309	6309

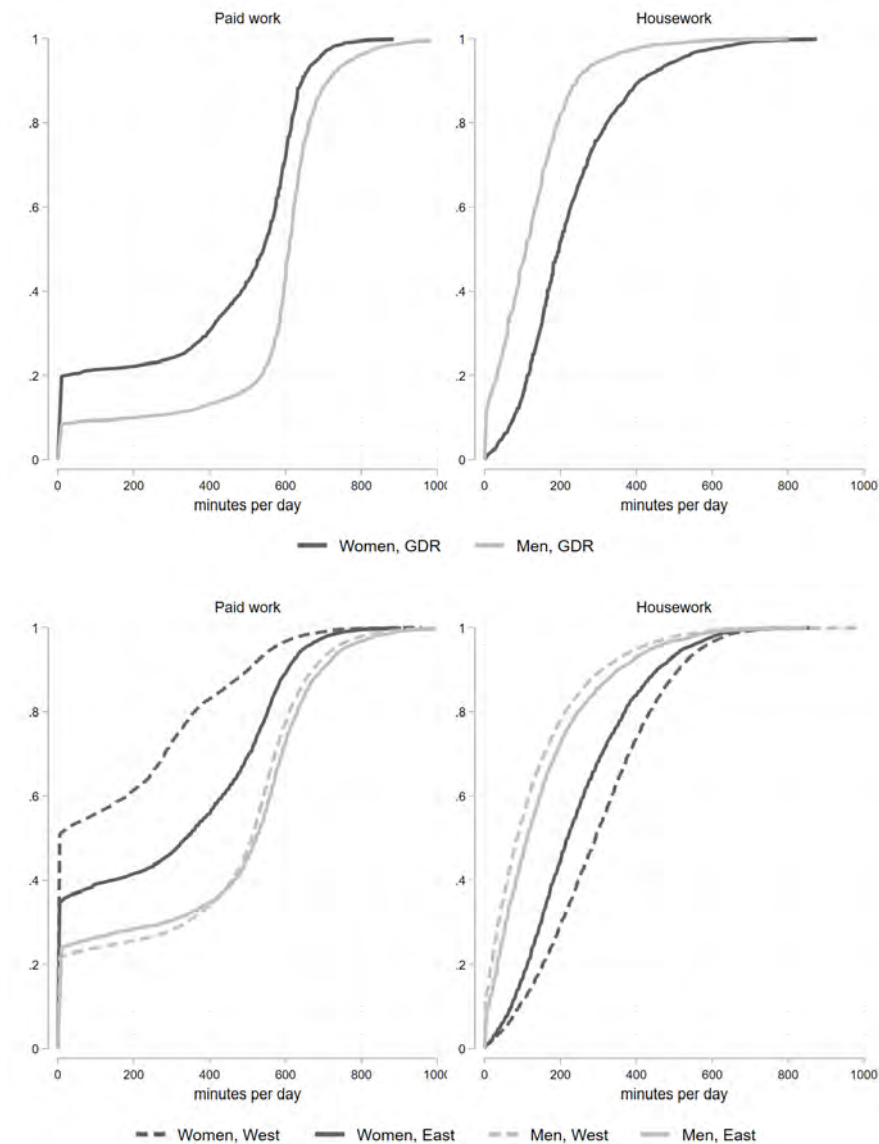
Notes: Table shows summary statistics of the time-use survey of the GDR and of the 1991/92 wave of the German time-use survey, separately for East and West Germany. Sources: GDR time budget study (1985/90) and German time budget study (1991/92).

6.4 Results

We start our analysis by illustrating differences in gender inequalities of time allocation between the GDR or East Germany, and West Germany. Figure 6.2 displays cumulative density functions of paid work and housework in the GDR, and East and West Germany, by gender. In the GDR only about 20% of women and less than 10% of men did not work on the weekday recorded. While almost all women do at least some housework, about 10% of men did not spend any time on housework. The lower panel of Figure 6.2 shows that despite the lower average in paid work, the distributional gender differences in the GDR in 1985/90 and East Germany 1991/92 were relatively similar. In West Germany, we see a much more gender-traditional distribution of paid and housework

work. Overall, women in East and West Germany differ strongly in their distributions, while men from East and West Germany look quite similar. The gender gaps in paid work and housework are substantially larger in West Germany than in East Germany, indicating that the GDR (and later East Germany) was more gender equitable. The lower panel in Table 6.1 presents average time use in minutes on weekdays.

Figure 6.2: Cumulative distributions of paid work and housework



Note: The figure plots cumulative density functions of paid and housework work among women and men. The upper panel uses the GDR time budget study (1985/90), the lower panel the German time-use survey (1991/92), separately for East (solid) and West Germany (dashed). Sources: GDR time budget study (1985/90) and German time budget study (1991/92).

6.4.1 Heterogeneity by household type

Figure 6.1 showed a strong inverse relationship between the female housework and paid work share using multinational data. In a next step, we analyse household heterogeneity by replicating this figure at the within-country household-level using data from the GDR. We define three types of households distinguished by the female share of market work; 1 $[0, 0.35)$, 2 $[0.35, 0.65)$ and 3 $[0.65, 1]$. Type 1 is a (main) male breadwinner household, whereas the second type is a dual-earner type / equal work household type. As laid out in section 6.2, the typical household in the GDR was of the second type. Type 3 households, (main) female breadwinner, remain the exception. We restrict this analysis to households where at least one partner is working full-time (≥ 7 hours) which are the most relevant cases in our context.¹²

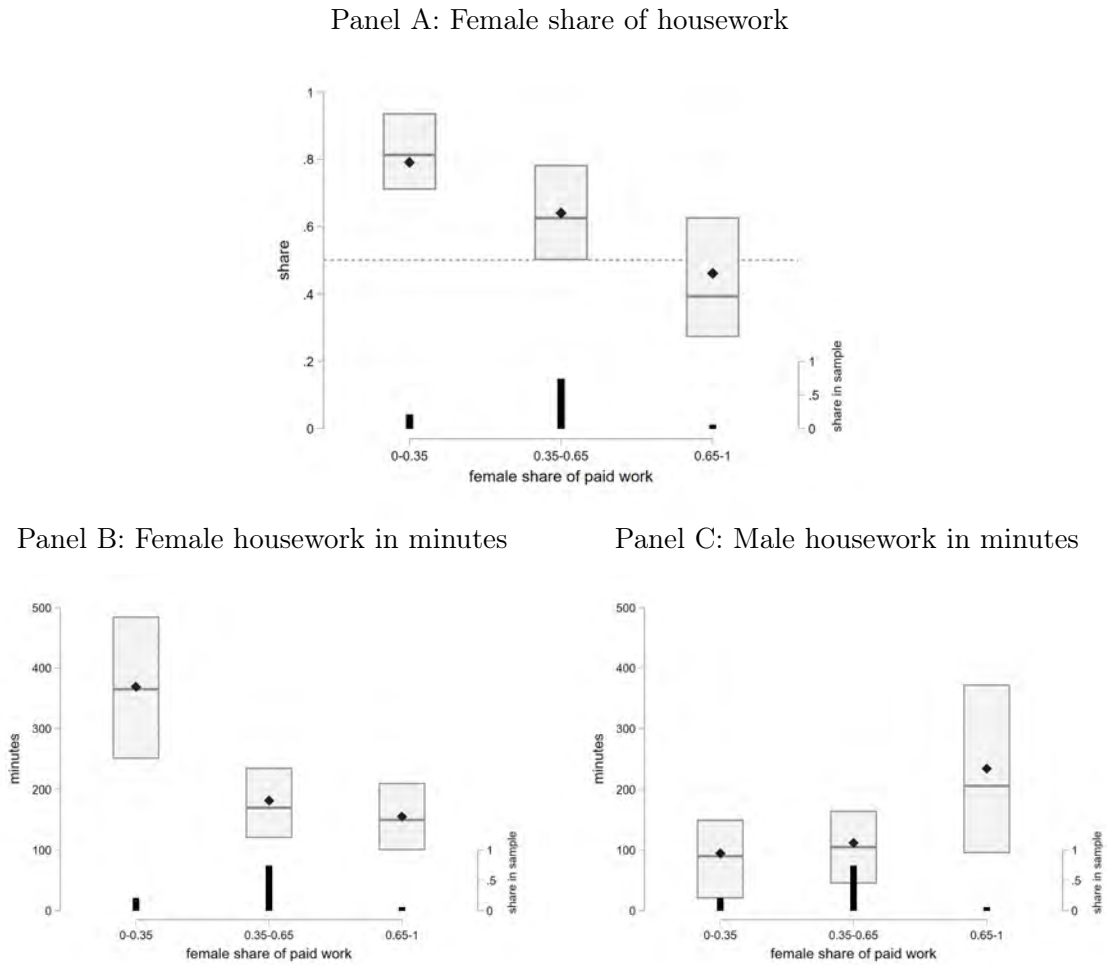
Figure 6.3, Panel A, shows the relation between the female share of housework and female share of paid work distinguished by household types. The range plots indicate 25th, 50th and 75th percentiles of the distribution of share of housework (diamonds show mean values). Panel A illustrates that the cross-country negative relation between household and market work shares also replicates across households of the GDR. Second, the lower solid black bars (right-hand Y -axis) shows the prevalence of the dual-earner household. 74 percent of households are the dual-earner type, only 20 percent male breadwinner, and a residual of five percent female breadwinner.

We next split the shares in housework into minutes contributed by females and males (Panels B and C), across our three types of households. The graphs now show the *total minutes* of housework of females and males. Consider the case of the male-breadwinner household. Here, females spend almost four times as much time doing housework compared to the working males: about 380 vs. 100 minutes. These figures underlay the share of almost 0.8 for the male-breadwinner household shown in Panel A.

Now, consider what happens when we move to the dual-earner household: Notably, females reduce their time spent doing housework substantially to below 200 minutes, halving their time spend in this activity. On the other hand, males in dual earner households still contribute about 100 minutes to housework. In other words, from the perspective of the male, it makes little difference if the female is working full-time, or few hours or not at all. But this is only one side of the coin: In the (rare) case of the female breadwinner household, males spend over 200 minutes on housework, whereas the full-time working females spend about 160 minutes. Comparing this to the dual-

¹²This excludes cases where both partners work few hours (27% of households) and a shift between different types of households can occur by small changes in one's working time, which only provides limited insight in this context.

Figure 6.3: Female share of housework, female and male minutes of housework by share of paid work in the GDR (1985/90)



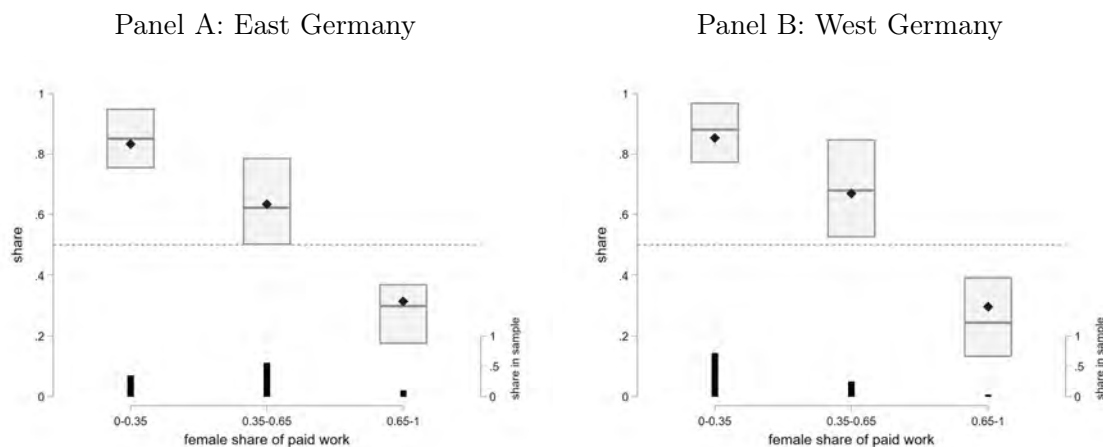
Notes: Panel A plots the females within-household share of housework against three intervals of share of paid work. Panels B and C show the female and male housework in minutes against intervals of share of paid work. Diamonds indicate the mean values, range plots show 25th, 50th and 75th percentiles of the distribution. Source: GDR time budget study (1985/90).

earner household, the adjustment only takes place by the partner who is changing their own labour supply.

Taken together, the gender-split of the inverse relation between household shares in time spent on the labour market and time spend doing housework suggests that the overall negative relation is almost entirely driven by individual time-constraints. Changes in the partner's work arrangement hardly affect own decisions, at least in this cross-sectional comparison across three stylised types of households.

We now repeat this analysis to compare East and West Germany where on average East Germany was shown to have a more equal distribution of market work and house-

Figure 6.4: Shares of housework by share of paid work in East and West Germany (1991/92)



Notes: Figure plots the males and females within household share of housework against three intervals of share of paid work. Source: German time budget study (1991/92).

work (Figure 6.2 and Table 6.1). To do this, we show results based on the 1991/92 time-use survey for reunified Germany. Figure 6.4 replicates Figure 6.3 for East and West Germany. At first sight, the two regions from Germany look remarkably similar with a comparable downward relationship between shares in market and household work. This may be counterintuitive at first as East Germany is commonly portrayed to be more gender egalitarian in many respects. However, the black bars indicating the distribution of household types reveal the underlying reason for this unexpected similarity in the graphs; the East German distribution resembles the one from the GDR strongly, whereas in West Germany the (main) male breadwinner type is most common. 72 percent of households belong to type 1 and only 25 percent are dual earner households. Once this differential selection into employment is taken into account within this simplistic framework, relative housework contributions differ little between East and West Germany. Appendix Figure 6.A.1 shows male and female contributions in minutes with a similar conclusion.

Taken together, the lack of reaction of individuals to their partner's changes in labour supply to market work, and the different incidence of household types, suggests that while large gaps in housework existed between the GDR (or East Germany following reunification) and the West, these need not necessarily reflect different behaviour at the level of the household, but are solely due to different selection into market work.

6.4.2 Decomposition of the gender gap in housework

To analyse East-West differences in a more structured way and including all household types, we now turn to a decomposition of the housework gap between East and West Germany. Again we use the German time-use survey from 1991/92, i.e. shortly after reunification. This has the advantage that while norms are arguably still strongly influenced by the differing environments individuals were exposed to during the German division, the survey is conducted in a uniform fashion, alleviating concerns about different survey designs or sample selections (which may be an issue when comparing the GDR data to West Germany). The overall gender housework gap varies strongly between East and West Germany (see Table 6.1 for summary statistics of the time-use data). While women perform 209 minutes more of housework than men on weekdays in West Germany, the difference is ‘only’ 102 minutes in East Germany.

To elicit to what degree the East-West difference of 107 minutes is due to observed factors, we estimate regressions controlling for important individual and household characteristics and, most importantly, male and female time spent in market work. If after controlling for these factors a large residual East-West difference in the housework gap remains, then this is strongly suggestive of other unobserved differences, such as different norms regarding gender roles in the household, which play an important role independent of norms regarding those concerning the labour market.

We use the conditional decomposition developed by Gelbach (2016) to analyse the role that different covariates play in explaining the East-West difference. In a first step, the housework gap is regressed on an East dummy. In a second step, the full model using all explanatory factors is estimated. Taking into account both the correlation between the dependent variables and the outcome variable as well as the correlation between the regions (East or West) and the dependent variables, the conditional decomposition from Gelbach (2016) provides consistent estimates on the role of each covariate in moving the East dummy from the baseline to the full model.

Results are presented in Table 6.2. Column 1 shows the raw difference, columns 2-5 contain the full model where we control for female and male market work in different ways (fixed effects for equally large categories of paid work or treating it as continuous). The East dummy is strongly reduced from 107 minutes to 11-16 minutes depending on the specification. I.e., a large share of the East-West gap can be explained by the covariates. Looking at the contributions of the different groups of explanatory variables, it is apparent that basic individual and household-level controls as well as household income have only a small impact. Depending on the specification, female work is responsible for 82-85 percent of the reduction of the East dummy. In contrast, male work, if anything, has only a minor effect. This supports the notion, in line with the

Table 6.2: Decomposition of the housework gap - East and West Germany (1991/92)

Dependent variable	Female-male housework gap (West mean: 209 minutes)				
	(1)	(2)	(3)	(4)	(5)
East dummy	-107.337*** (5.435)	-12.291** (4.611)	-11.231* (4.551)	-15.700*** (4.461)	-12.572** (4.478)
<i>Covariates:</i>					
Basic controls		-7.084*** (1.117)	-7.145*** (1.127)	-7.049*** (1.111)	-7.347*** (1.133)
Household income (5 categories)		-3.922 (2.115)	-4.160* (2.094)	-4.675* (2.088)	-4.278* (2.083)
Female paid work		-77.829*** (3.578)	-80.003*** (3.647)	-77.963*** (3.587)	-78.589*** (3.601)
Male paid work		-6.211* (3.041)	-4.798 (3.098)	-1.951 (3.037)	-4.551 (3.103)
Specification of female / male work		5 FEs	10 FEs	linear	lin. & sq.
Share coef. movement due to female work		.819	.832	.851	.829
Observations		6,380	6,380	6,380	6,380

Notes: Table shows a decomposition of the female-male housework gap following Gelbach (2016). Lower rows show the contribution of the groups of explanatory variables in moving the East dummy. Basic controls: Children under 10 in household (0/1), household size, age, partner's age, education dummies. Robust standard errors in parentheses. Source: German Time-Use Survey

visual evidence in Figure 6.4 (share of housework) and Appendix Figure 6.A.1 (minutes of housework), that it is not relative contributions to paid work that determines housework, but to a large extent only women's own contribution. East-West differences in the housework gap are thus mostly due to different selection into employment of women. The more gender egalitarian division of household tasks within the GDR that is documented by overall averages should not be mistaken for evidence for more gender egalitarian behaviour at the level of the household.¹³

The analysis is also repeated with later waves (2001/02 and 2012/13) of the German time use survey, shown in Appendix Table 6.A.5.¹⁴ The average gap reduces strongly over time and so do unconditional East-West differences. Once female work is controlled for, the differences are small and no significant differences between the waves can be observed. The result suggests that later spill-overs of differences between the GDR and FRG into the domain of housework are unlikely. Moreover, they support the conclusion that once individual work is accounted for, differences between regions are insignificant even way beyond the period of German reunification, where many institutions changed.

¹³In Appendix Table 6.A.3 the same table is shown, but with the dependent variable specified as the female housework share. Reassuringly, the results are very similar. Appendix Table 6.A.4 shows results restricted to the subcategories *cooking, cleaning, and shopping*. These "routine" domains of housework are the most time-consuming and are "less optional and less able to be postponed" (Coltrane, 2000). Women are commonly specialised in routine housework (Borra et al., 2021). Overall patterns are very similar.

¹⁴As results from different specifications of female and male work in Table 6.2 were in the same ballpark, we proceed using the specification with 5 categories of paid work as dummy variables.

6.4.3 The role of children

We study the role of children in more detail as institutional differences in child care availability between the GDR and the FRG, particularly for children under the age of 3, which persisted after reunification, are a potential explanation for small remaining East-West differences in conditional housework gender gaps in Table 6.2. Having young children in child care instead of home care might reduce the amount of housework needed, which then causes West German women to work more in the household than East German women, even conditional on employment.

Results separately for couples with children under the age of 10 and couples with only older or no children are presented in Table 6.3. Columns 1-4 contain estimates for the female-male housework gap. In families with no young children (columns 1-2), controlling for the covariates leads the initial large East dummy to be small and statistically insignificant. The coefficient movement of the East dummy is almost entirely driven by female work. As institutional differences in child care provision (and afternoon care for primary school children) between East and West are irrelevant for this group, this is arguably the cleanest comparison for housework norms conditional on observed covariates.

Table 6.3: Decomposition of the housework gap by children - East and West Germany (1991/92)

Dependent variable	Female-male housework gap				Female-male domestic gap	
	206 minutes		213 minutes		307 minutes	
West mean:						
Sample:	No children u10		Children u10			
	(1)	(2)	(3)	(4)	(5)	(6)
East dummy	-90.829*** (7.548)	-5.008 (5.889)	-131.369*** (7.495)	-37.557*** (7.394)	-163.433*** (9.520)	-40.006*** (8.338)
<i>Covariates:</i>						
Basic controls		-5.183*** (1.130)		-8.181*** (1.956)		-0.436 (1.808)
Household income (5 categories)		-5.448 (2.944)		0.919 (2.898)		2.771 (3.332)
Female work		-69.168*** (5.063)		-79.720*** (4.824)		-117.347*** (6.657)
Male work		-6.022 (4.376)		-6.829 (3.857)		-8.415 (5.071)
Specification of female / male work		5 FEs		5 FEs		5 FEs
Share coef. movement due to female work		.806		.85		.951
Observations	3,868	3,868	2,512	2,512	2,512	2,512

Notes: Table shows a decomposition of the female-male housework gap and domestic work gap (housework + child care) following Gelbach (2016). Samples are split by whether children under 10 live in households. Lower rows show the contribution of the groups of explanatory variables in moving the East dummy. Basic controls: household size, age, partner's age, education dummies. Robust standard errors in parentheses. Source: German Time-Use Survey

Columns 3-4 contain the estimates for households with children. While the East dummy is strongly reduced by the covariates, the larger remaining gap of 38 minutes (just below 30%) highlights the relieving effect of the East German child care infrastructure for mothers.

In an additional step (columns 5-6), we change the dependent variable to the female-male *domestic* gap, which in addition to housework also includes time spent on child care activities. The covariates, again especially selection into employment for women, strongly reduce the East dummy, but it remains at a similar magnitude than for the housework gap (column 4). All in all, the distinction by children stresses that overall East-West differences in the gender-housework gap are small once employment is considered, but they remain stronger when children are part of the equation.

6.5 Conclusion

We use five waves of time-use data from the GDR, and East and West Germany to examine the relation between time spent on the labour market and time spent doing housework. We document that women in the GDR, and later in the East, spend more time in the labour market and less time doing housework. However, the gap in time spent doing household tasks between East and West Germany can almost fully be explained by individual time constraints—if women are working more, they mechanically reduce the time spent doing housework, because the day only has 24 hours. In contrast, we find little evidence that labour supply of the partner matters. As a result, the housework allocation in both Germanies is very similar, once own working time is accounted for. In this regard, Germany was always united.

This finding has implications for models of household-bargaining that typically assume that earnings potentials matter for the time-allocation to unpleasant tasks. We find that the partner-reactions to institutionally-induced changes in earnings potential matter little for within-household time-allocation. One explanation for this is that gender norms regarding household work evolve independent of gender norms in the labour market, and in the context studied are strong enough to cover-up any potential effects of labour market returns on bargaining power. The position of the data points studied by us in relation to other countries (Figure 6.1) suggests this is not merely a German phenomenon.

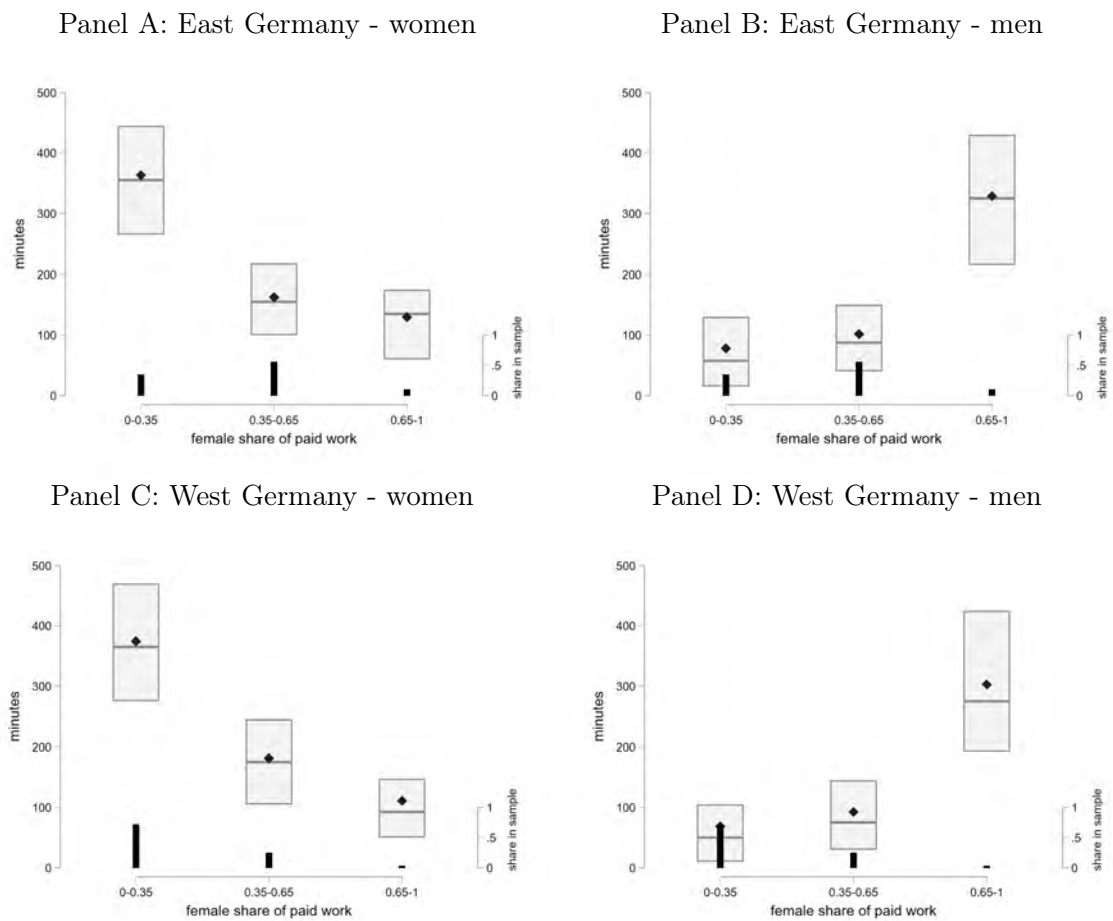
We conclude that labour market policies targeted at increasing women's time spent working must not lead to changes in the household allocation process beyond a mechanical effect. If motivated by ideas of female empowerment and gender equality, such policies could even backfire and push women into the second shift, where women have to squeeze own working time and household tasks into their time budget. We

believe that investigating such, potentially negative, implications for female utility is an exciting route for future research.

6.A Appendix

6.A.1 Figures

Figure 6.A.1: Minutes of housework by gender and share of paid work in East and West Germany (1991/92)



Notes: Figure plots male and female minutes of housework against three intervals of share of paid work. Data from the German time budget study (1991/92).

6.A.2 Tables

Table 6.A.1: Countries and surveys used from the Multinational Time Use Study

Country	Survey year	Observations
Austria	1992	3,571
Bulgaria	2001	1,042
Canada	1992	3,571
Denmark	2001	1,002
Finland	1999	1,055
	2009	729
France	1998	2,275
	2009	2,602
Israel	1991	534
Italy	1989	2,582
	2002	2,716
	2008	2,292
Slovenia	2000	957
South Africa	2000	684
South Korea	1999	11,405
	2009	5,240
Spain	2002	5,010
	2009	2,043
United Kingdom	1974	2,867
	1984	461
	1987	1,971
	2000	1,864
	2014	1,138

Notes: Table shows the list of countries, surveys and number of observations that are used from the Multinational Time-Use Study. Samples are restricted to different-sex couples aged 18-65. Observation numbers are at the couple-level. We use all countries and survey waves from the Multinational Time Use Study with a household-level sampling design and the existence of household identifiers in the data. Finally, we require to have at least 400 couple-day-level observations.

Table 6.A.2: Time-use data: detailed activities (minutes per weekday)

	GDR (85 and 90)		East Germany (91/92)		West Germany (91/92)	
	Women	Men	Women	Men	Women	Men
Paid work	429.27	556.55	325.87	462.42	157.64	467.21
Working	364.00	471.86	266.82	396.24	131.10	403.14
Work-related (breaks, travel time etc.)	64.09	82.20	42.76	54.54	17.90	57.45
School / studies	1.19	2.49	16.29	11.64	8.64	6.63
Housework	223.78	121.89	250.55	148.86	325.55	116.53
Cooking	59.50	13.07	85.29	24.92	106.92	17.24
Cleaning	79.41	9.44	78.78	13.91	119.26	10.59
Fixing and building things	14.76	29.43	8.10	40.62	8.29	32.41
Shopping	37.76	19.40	25.39	14.63	32.16	11.99
Gardening	21.71	43.19	19.87	26.44	23.86	20.67
Other housework	10.64	7.36	33.12	28.34	35.05	23.62
Care for others	45.14	13.15	60.58	25.48	77.63	22.99
Child care	41.59	11.80	47.87	18.23	59.10	16.91
Care for adults	3.55	1.35	2.25	1.00	5.38	1.63
Leisure	157.98	194.18	184.12	203.60	229.91	221.28
Cultural activities	16.00	21.57	2.30	3.65	7.71	7.04
Sports (active and passive)	12.32	13.07	11.83	14.66	21.32	22.06
Media consumption	90.38	120.23	109.34	132.90	111.95	129.31
Social contacts	24.35	22.79	51.35	43.62	74.88	50.17
Other leisure	24.78	24.10	9.29	8.77	14.04	12.71
Observations	2328	2328	1673	1673	4707	4707

Notes: Table shows fine-grained activities that are contributing to the broader categories the analyses build on.

Table 6.A.3: Decomposition of the female housework share - East and West Germany (1991/92)

Dependent variable	Female-male housework gap (West mean: 0.751)				
	(1)	(2)	(3)	(4)	(5)
East dummy	-0.102*** (0.007)	-0.013* (0.006)	-0.013* (0.006)	-0.019** (0.006)	-0.014* (0.006)
<i>Covariates:</i>					
Basic controls		-0.005*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)
Household income (5 categories)		-0.010*** (0.003)	-0.010*** (0.003)	-0.010*** (0.003)	-0.010*** (0.003)
Female work		-0.071*** (0.004)	-0.074*** (0.004)	-0.065*** (0.003)	-0.073*** (0.004)
Male work		-0.003 (0.004)	-0.000 (0.004)	-0.003 (0.004)	0.000 (0.004)
Specification of female / male work		5 FEs	10 FEs	linear	lin. & sq.
Share coef. movement due to female work		.794	.826	.783	.826
Observations		6,372	6,372	6,372	6,372

Notes: Table shows a decomposition of the (log) female-male housework gap following Gelbach (2016). Lower rows show the contribution of the groups of explanatory variables in moving the East dummy. Basic controls: Children under 10 in household (0/1), household size, age, partner's age, education dummies. Robust standard errors in parentheses. Source: German Time-Use Survey

Table 6.A.4: Decomposition of the housework gap (cooking, cleaning, shopping) - East and West Germany (1991/92)

Dependent variable	Female-male housework gap (West mean: 219 minutes)				
	(1)	(2)	(3)	(4)	(5)
East dummy	-82.526*** (3.968)	-12.993*** (3.836)	-11.800** (3.809)	-14.425*** (3.740)	-12.384*** (3.741)
<i>Covariates:</i>					
Basic controls		-7.100*** (1.019)	-7.065*** (1.022)	-7.011*** (1.016)	-7.202*** (1.030)
Household income		0.656 (1.766)	0.450 (1.752)	0.102 (1.749)	0.355 (1.746)
Female paid work		-60.064*** (2.780)	-61.610*** (2.828)	-60.355*** (2.796)	-60.821*** (2.793)
Male paid work		-3.025* (1.346)	-2.501 (1.373)	-0.837 (1.303)	-2.473 (1.366)
Specification of female / male work		5 FEs	10 FEs	linear	lin. & sq.
Share coef. movement due to female work		.864	.871	.886	.867
Observations		6,380	6,380	6,380	6,380

Notes: Table shows a decomposition of the female-male housework gap following Gelbach (2016). Lower rows show the contribution of the groups of explanatory variables in moving the East dummy. Basic controls: Children under 10 in household (0/1), household size, age, partner's age, education dummies. The housework definition in this table is more narrow than the one shown in Table 6.2. See Table 6.A.2 for an overview of subcategories of household. Robust standard errors in parentheses. Source: German Time-Use Survey

Table 6.A.5: Decomposition of the housework gap - East and West Germany (2001/02, 2012/13 and pooled)

Dependent variable	Female-male housework gap					
	165 minutes		126 minutes		172 minutes	
West mean:	2001/02		2012/13		Pooled	
Survey wave:	(1)	(2)	(3)	(4)	(5)	(6)
East dummy	-71.622*** (6.857)	-9.184 (5.107)	-52.700*** (6.550)	-11.783* (4.903)	-83.523*** (3.611)	-13.565*** (2.744)
<i>Covariates:</i>						
Basic controls		-6.054*** (1.453)		-5.201*** (1.409)		-6.382*** (0.719)
Household income (5 categories)		1.082 (0.977)		0.324 (0.958)		0.139 (0.723)
Female paid work		-50.926*** (4.381)		-33.370*** (3.910)		-59.131*** (2.293)
Male paid work		-6.539 (4.541)		-2.670 (3.965)		-1.998 (2.189)
Survey wave						2.009*** (0.394)
Specification of female / male work		5 FEs		5 FEs		5 FEs
Share coef. movement due to female work		.816		.816		.905
Observations	5,313	5,313	4,039	4,039	15,732	15,732

Notes: Table shows a decomposition of the female-male housework gap following Gelbach (2016) by survey wave. Lower rows show the contribution of the groups of explanatory variables in moving the East dummy. Basic controls: household size, age, partner's age, education dummies. Robust standard errors in parentheses. Source: German Time-Use Survey

CHAPTER 7

Culture, Children and Couple Gender Inequality¹

7.1 Introduction

Women’s labour force participation has increased strongly across high-income countries in past decades and gender inequality has been reduced on many domains. Yet, despite this progress, women tend to work fewer hours than men, gender wage gaps remain substantial, and in few couples is the woman outearning her partner. Those inequalities are remarkably persistent (Olivetti and Petrongolo, 2016). As more women than men in high-income countries hold college degrees (Kleven and Landais, 2017), classic human capital models fail to account for persisting gender inequalities. The literature has identified children as a main source of remaining gaps (Córtes and Pan, 2020; Kleven et al., 2019b), but uneven labour market responses to becoming a parent—i.e. employment interruptions with limited recovery are commonly only observed for mothers—are not per se deterministic. Some institutional features, such as more generous parental leave allowances for mothers, favour longer leave taking by mothers and, more generally, main breadwinner models (e.g. joint tax filing for spouses). More recently, the role of culture in determining maternal employment has received increased attention with a particular focus on intergenerational transmittance (e.g. Fernández, 2007; Fernández and Fogli, 2009). Giuliano (2021) provides a comprehensive overview of the literature on gender and culture.²

In this paper, I examine how culture impacts within-couple gender inequality. Using a long running household panel (SOEP), I compare child penalties using event study estimates between couples socialised in a more gender-egalitarian culture to those in a more gender-traditional culture but living in the same country. For this, I exploit

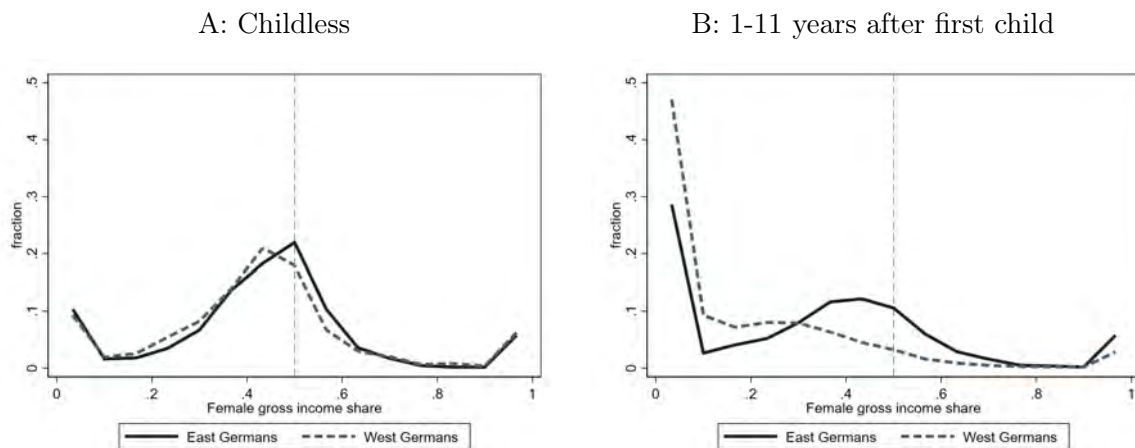
¹I am grateful to Jan Berkes, Ludovica Gambaro, Jan Marcus, Anna Raute, Viola Salvestrini, Almudena Sevilla, C. Katharina Spiess, Guo Xu and audiences at the Queen Mary University of London, DIW Berlin, BSE Applied Micro Workshop for helpful comments.

²In line with Giuliano (2021), I borrow the definition of culture proposed by Guiso et al. (2006) as a set of beliefs and values held by groups that are transmitted over generations.

the unique setting of Germany’s division and reunification, where couples growing up in the German Democratic Republic (GDR) were exposed to more gender egalitarian policies³ and norms than those in the Federal Republic of Germany (FRG),⁴ especially regarding maternal employment.

Figure 7.1 plots the distribution of the female share of household income in East and West German couples.⁵ For childless couples (Panel A) the income distributions are almost identical with a modal share of just below half, i.e. pre-birth gender differences exist but they are quite small. Additionally, with 34% (East) and 29% (West), a non-negligible share of women has higher earnings than their partner. In contrast, East-West differences are striking for parents with young children (Panel B); while the female income share has plummeted for both, the distribution for West German couples is substantially more right-skewed with a large share of mothers having exited the labour force or working few hours. This is also reflected in the shares of couples with young children having a main female earner, which are 23% (East) and 8% (West). Barth et al. (2020) also show that differences in maternal full-time employment have decreased following reunification but no further convergence has occurred since the early 2000s.

Figure 7.1: Female income shares in East and West German couples



Notes: Income share based on gross monthly labour income of both partners. Distribution calculated in 15 bins of equal width. In Panel A the age range of women is restricted to be between the 5th and 95th percentile of Panel B (25 to 44). Sample covers 1990-2019. Source: SOEP v36

³E.g. to facilitate maternal employment day care provision was universal in the GDR and Article 18 of the 1949 constitution already stated “equal pay for equal work” (Trappe, 1996). Also since the early 1950s women could freely decide on their employment, in the FRG this was only the case since 1977 (Lippmann et al., 2020).

⁴For simplicity, throughout this paper I use GDR and FRG when referring to the two German states before reunification, and to East Germany and West Germany after reunification.

⁵East and West German couples are defined by their location in 1989, see section 7.3.

Several papers have shown that gender norms of parents map into those of their children. Farré and Vella (2013) examine intergenerational correlations in gender attitudes and find that mothers' attitudes have a strong effect on those of their children when these are adults. Kleven et al. (2019b) estimate the intergenerational transmission of child penalties and argue that parents' gender norms form their daughters' norms during childhood. In their ground-breaking work, Fernández et al. (2004) document that wives of men who grew up with their mothers working are more likely to be in the labour force themselves.⁶ As a result, different socialisation in the GDR or FRG may have long-lasting impacts on child penalties and gender inequality.

I first estimate event-studies with household panel data and find that *labour market inequality* due to children is substantially stronger in West German couples with a negative long-run effect on the female earnings share of 26.9 percentage points (pp), 11.4 pp larger than in East German couples. I then show that inequality in *unpaid domestic work*, housework and child care, similarly increases strongly upon the arrival of children, with effects again being more moderate in East German couples (about 8 pp lower for both outcomes). Contrary to labour market outcomes, inequality in housework was already pronounced before couples have children. A summarising specialisation index reveals that a gender-traditional re-orientation is more than twice as strong in West German couples.

I add to this analysis by using time-use diary data from the GDR and reunified Germany to look at gender inequality in time-use by children in more detail. Micro data from the GDR regime is scarce and the newly digitised time-use data thus allows to examine gender inequality in a regime with one of the world's highest female labour force participation rates. The data show that gender-specific differences in working time in the GDR were much smaller than in post-reunification West Germany (and to a lesser degree, East Germany) and that the child penalty for women was smaller. Inequalities in domestic work on the other hand were also strong in the GDR and women were almost solely responsible for child rearing. As the time-use data is cross-sectional, this analysis contrasts couples with and without children, but of a similar age range and controlling for important observable characteristics.

In a final step, using another household panel (pairfam), I analyse differences in attitudes towards maternal employment and how those attitudes are affected by children. East Germans favour longer working hours for mothers at all child ages, except in the first year when the labour market effect of children is also similar. East Germans are less likely to agree that women should prioritise family over career and that a working mother is harmful for children under 6. Using event study estimates, I then show that

⁶Schmitz and Spiess (2021) identify the same mechanism in West Germany.

children lead to more gender traditional attitudes for East and West Germans, with suggestive evidence for a slight convergence of attitudes.

A main contribution of this paper is that it estimates child penalties through the lens of social norms. A growing literature has estimated child penalties in recent years in different countries and settings (e.g. Angelov et al., 2016; Bertrand et al., 2010; Córtes and Pan, 2020; Kleven et al., 2019a,b; Kuziemko et al., 2018). A consistent finding is that mothers' labour market trajectories are strongly affected in the short-run without full recovery. Effects on fathers tend to be small. In their paper on child penalties in Swedish couples, Angelov et al. (2016) focus the heterogeneity analysis on relative educational attainment and find that the within-couple gap disappears four years after birth only when mothers have a substantial educational advantage.⁷ Kleven et al. (2019b) study child penalties in Denmark, documenting underlying mechanisms in detail (e.g. selection into more child-friendly occupations after birth) and showing the transmittance of child penalties across generations; child penalties are closely linked to the labour supply of maternal grandparents. Kleven et al. (2019a) conduct a cross-country analysis of child penalties and show that these are much lower in Scandinavian countries compared to the US, UK, Austria and Germany, and that penalties are closely linked to stated gender norms. Building on the two latter papers, this paper estimates child penalties *within* one country, where during the German division individuals were exposed to different policies and gender norms.

A recently emerging literature has compared the impact of children on East and West German mothers. Collischon et al. (2020) contrast child penalties for employment, working hours and hourly wages. Using rich administrative data, Boelmann et al. (2021) address a similar question, but they take several steps to convincingly control for potentially confounding factors and explore further mechanisms.⁸ I add to those papers in several dimensions; by taking a holistic view of children and gender inequality, besides looking at labour market outcomes, I additionally examine differences in time allocation in the household to non-market work (housework and child care) and to what extent the arrival of children induces a change in attitudes. As an additional contribution, I use time-use data from the GDR, giving a rare insight into gender

⁷In contrast, Córtes and Pan (2020) and Kleven et al. (2018) find no evidence for strong heterogeneities by relative education in the US and Denmark, respectively.

⁸Boelmann et al. (2021) first document persistent differences within cross-border labour markets, second they show, by looking at migrating mothers, that East Germans in the West keep their norms whereas West Germans in East Germany adjust to local gender norms, and, finally, they document that West German mothers with a high inflow of East Germans in their firm adjust their post-birth return behaviour in the direction of East German mothers. The same local learning mechanism has also been found by Schmitz and Weinhardt (2019) who take a macro-perspective by examining how West German women's labour force participation changes when their counties have experienced a high inflow of East Germans in the years following reunification.

inequality in a state socialist regime. Combined with time-use data from reunified Germany, I can compare inequalities in the GDR with those in East and West Germany in a consistent framework.

This paper also takes a couple-perspective, which is particularly valuable in this context. When differences in child penalties between groups are of key interest, the couple-perspective automatically controls for potential contextual confounders. In the case of East vs. West Germany, whereas the institutional framework, e.g. parental leave or the tax system are identical,⁹ labour market conditions and day care supply are factors with regional discrepancies and aspects that could impact child penalties differently. On the couple-level, these are automatically accounted for. Finally, because children have been found to be by far the biggest source of residual gender inequality in earnings in the 2010s,¹⁰ by looking at children and their effect on couple gender inequality, the lens is put on by far the most important aspect of overall gender inequality in the labour market and direct inference can be drawn on gender gaps of parents (Angelov et al., 2016).¹¹

Several influential papers on gender inequality have also taken a couple-perspective. Most notably, in their seminal work, Bertrand et al. (2015) look at gender identity norms and relative income within married couples in the US, identifying strong aversion to a situation of the wife outearning her husband. Building on this, Lippmann et al. (2020) compare East and West German couples and find that exposure to more gender equal institutions has indeed *undone gender norms*, as East German women can have higher earnings within a couple without increasing housework (see West and Zimmerman, 1987, for the "doing gender" hypothesis) or risking their marriage. In West German couples those consequences of traditional gender norms are still prevalent.¹² However, Lippmann et al. (2020) do not explicitly consider the role of children for gender inequality and only control for the presence of children in their estimation. As children are the main source of differential within-couple gender inequality between

⁹An exception of the same institutional environment is the different upper earnings limit for statutory pension insurance, which as of 2004 (the median year of the analysis) at 5,150 Euros per month was 18% higher in West Germany.

¹⁰Two-thirds of gender inequality in the US and 80% in Denmark is child-induced, see Córtes and Pan (2020) and Kleven et al. (2019b) respectively.

¹¹An additional aspect that makes the couple-perspective more relevant is the increasing role of fathers in child rearing. Gimenez-Nadal and Sevilla (2012) show that fathers' child care involvement (and other unpaid work such as housework) has increased substantially over the past decades across high-income countries—albeit to still much lower levels than that of mothers. While studies commonly find that fathers' labour market outcomes are, if anything, only marginally affected by the arrival of children (e.g. Bertrand et al., 2010; Kleven et al., 2019b), fathers may react in their involvement in child care and in other domains of non-market work such as household chores.

¹²Sprengholz et al. (2020) investigate a similar question with the same data using annual rather than monthly earnings measures, but are unable to confirm this finding.

East and West Germans (see Figure 7.1), estimating child penalties sheds light on the magnitude of children in explaining this.

This paper also contributes to the sizeable literature examining long-run *effects*¹³ of exposure to the two German regimes on a wide range of outcomes. Papers studying gender-related attitudes have consistently found more gender-egalitarian views in East Germany with limited signs of convergence; this holds for the role of mothers in the labour market and in the family (Bauernschuster and Rainer, 2012), gender-specific work preferences (Beblo and Görge, 2018), importance of career success for women (Campa and Serafinelli, 2019), and attitudes about detrimental effects of maternal employment on children (Zoch, 2021).¹⁴ In line with those attitudes, a more even distribution in household tasks (Cooke, 2007) and female income share (Lippmann et al., 2020; Sprengholz et al., 2020) has been documented. I add the important dimension on how children impact differences labour market outcomes, domestic work, and attitudes.

The paper proceeds as follows. The next section discusses the historical context of the German division and reunification, section 7.3 describes the data sources used and outlines the empirical approach. Results are presented in section 7.4 followed by a battery of robustness checks in section 7.5. I conclude in section 7.6.

7.2 German division and reunification

After World War II, Germany was partitioned into four occupation zones. After increasing tensions in the post-war years, in May 1949 the Federal Republic of Germany (FRG) was formally established consisting of the three western zones, followed by the German Democratic Republic (GDR) in October 1949 consisting of the Soviet occupation zone. The two German states were to exist separately for 41 years.

The GDR and FRG followed very different paths when it came to policies regarding female employment and gender inequality (Trappe, 1996). The GDR—a socialist, de-facto one-party state—promoted a more gender egalitarian way, and both mothers and fathers in general worked full-time. This was actively stimulated by the GDR through the provision of a universal day care system and an obligation for both men and women

¹³Becker et al. (2020) have recently highlighted pre-existing differences between East and West Germany before the formal separation in 1949 as well as selective migration in the following years, due to which the German division cannot be treated as a clean *natural experiment* to study the long-run effects of communism / socialism (as many papers explicitly state). However, I do not claim to identify the effect of a political regime, but rather use the setting to compare child-induced gender inequality between regions with differing gender attitudes and histories of maternal employment.

¹⁴Other papers have, e.g., looked at differences in precautionary savings behaviour (Fuchs-Schündeln and Schündeln, 2005), preferences for redistribution (Alesina and Fuchs-Schündeln, 2007), or attitudes towards financial markets and investment behaviour (Laudenbach et al., 2020).

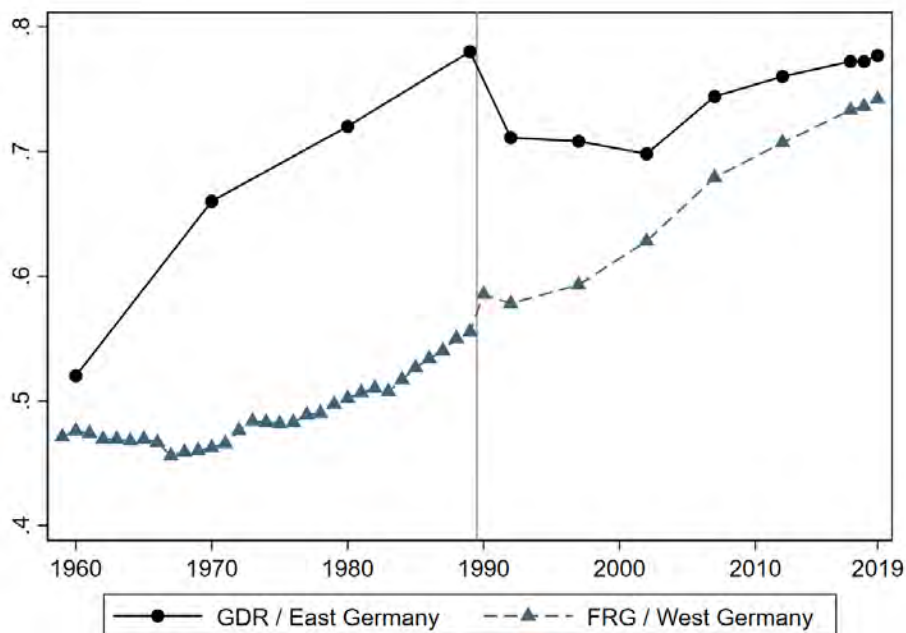
to be in employment (Beblo and G6rges, 2018).¹⁵ In contrast, the FRG was a market-based democracy with gender-conservative policies. Day care provision was limited, and the tax and transfer system encouraged a male breadwinner model (or a one-and-a-half male breadwinner model with the woman working part-time). A series of parental leave expansions in the 1970s and 1980s temporarily prolonged maternal leave, but long-run effects on labour market outcomes were limited (Sch6nberg and Ludsteck, 2014). Both before and after those reforms, a large share of mothers did not return to the labour market and, if so, mostly part-time. Differences in attitudes towards maternal employment were also pronounced as can be seen in derogatory nicknames *working* mothers were given in the FRG ("raven mothers") and *non-working* mothers in the GDR ("parasites", see Boelmann et al., 2021).

After increasing discontent in the GDR accompanied by mass demonstrations, the fall of the Berlin Wall in November 1989 finally led to reunification of the two German states in October 1990. In the direct aftermath large East-West migration streams began. In 1989 and 1990 alone, more than 800,000 East Germans migrated West, predominantly 18-30 years olds (Fuchs-Sch6ndeln and Sch6ndeln, 2009). In the reunification process the GDR was fully integrated into the FRG and adopted their policies, including the tax and transfer system, and parental leave legislation (since 1992 mothers had 36 months of employment protection and means-tested benefits of about 300 Euros for 24 months, Sch6nberg and Ludsteck, 2014). Yet some differences in the institutional environment remained, such as the larger day care availability in East Germany, a higher share working in the public sector and an overall weaker labour market (Rosenfeld et al., 2004).

Figure 7.2 shows female labour force participation rates for East and West Germany starting from 1959 to 2019. Differences were initially relative small, but the policies in the GDR led to a large increase in the following decades and reached 78% in 1989, among the highest rate in the world (Rosenfeld et al., 2004). On the other hand, participation in the FRG only increased slowly from the 1970s onward and before reunification female labour force participation was 22 pp lower than in the GDR. Despite an initial convergence in the years after reunification, differences have persisted over the past two decades. In line with differences in female labour force participation, research has also shown that attitudes towards maternal employment immediately following reunification were substantially more gender-egalitarian in East Germany (see e.g. Bauernschuster and Rainer, 2012).

¹⁵In 1976 a baby year was introduced for higher order births in the GDR and this was extended to all births in 1986 (Heisig and Zierow, 2021). During the baby year mothers received generous wage replacement, but commonly returned to employment thereafter. Fathers were in principle also eligible but rarely used it.

Figure 7.2: Female labour force participation



Notes: Figure shows female labour force participation for East Germany (GDR before 1990) and West Germany over time. The vertical line denotes the fall of the Berlin Wall. Sources: GDR statistical office (from Schmitz and Weinhardt, 2019), Destatis with Microcensus

7.3 Data and empirical approach

The main empirical analysis relies on the German Socio-Economic Panel (SOEP), a longitudinal household survey by the German Institute for Economic Research (DIW Berlin, Goebel et al., 2019). The survey started in 1984 in the FRG and added GDR households in 1990 before reunification was completed. Currently, SOEP contains about 15,000 households and 35,000 individuals per year. A wide range of topics are covered in the study, including labour market outcomes, attitudes, time-use, relationship details and socio-economic background characteristics. Being a panel study on the household-level, the data contain information from all household members aged 12 years and older. Importantly for my analysis, the survey asks where respondents had lived in 1989 (GDR, FRG or abroad), i.e. before reunification. As mobility between the GDR and FRG was strongly restricted, this variable indicates where respondents' parents grew up and where they themselves were socialised. I use the 1989 location to define East and West German couples.

7.3.1 Sample criteria and outcomes

This paper takes a couple-perspective on gender inequality and thus relies on the household structure of the survey. I focus on (becoming) parents to examine the role of children impacting gender inequality differentially in East and West German couples. In contrast to studies using administrative data (e.g. Angelov et al., 2016; Kleven et al., 2019b), imposing a balanced sample over a longer pre- and post-birth period would strongly reduce the sample. First, individuals from survey households may not always be covered from 3 years pre- to 6 years post-birth (this is the main sample window, whereby the upper limit is chosen to cover the usual age of school entry). Second, if for a couple full coverage is required, this implies that the couple must have formed a household before the window and not broken up until it ends, which would make the sample more selective, especially in the pre-birth period. The main results simply demand any couple observation in the event window, requiring a couple-level observations at least once before and after birth of the first child, similar to Córtes and Pan (2020), yields comparable results (see Appendix Table 7.A.1). Appendix Figure 7.A.1 shows how the observation window by household is distributed.

As this paper investigates gender inequality, same-sex couples are not considered in the analysis. Due to the large share of non-marital births in East Germany (58% vs. 27% in West Germany in 2009, see Klüsener and Goldstein, 2016), both married and non-married cohabiting couples are included (in contrast to the analyses of Lippmann et al., 2020; Sprengholz et al., 2020). A further requirement is that both partners have lived in the GDR or FRG in 1989. No further restriction is set on a migrant background. Due to the low share of mixed East-West couples (6.6%), the analysis focuses on single-origin couples. Overall I look at couples in working-age population (18-65), but the years surrounding the first-birth often impose a stronger restriction on the age range. While some papers examining earnings distribution of couples restrict their analysis to dual-earners couples (e.g. Bertrand et al., 2015; Lippmann et al., 2020), I keep observations where either partner has zero earnings, as especially mothers often (temporarily) drop out of the labour force in the years following birth and report zero earnings. To ensure comparability between households from the East and West, the sample is restricted to 1990 to 2019 where both are covered. Table 7.1 provides an overview of the number of observations by sample restriction.

The main labour market outcome is the share of female income of the household income within a couple. The income variable refers to gross labour income of the previous calendar month. As capital income is arguably to a lesser degree affected by gender norms in couples, this income component is not taken into account. As alternative measures of the income distribution in couples, results for *gaps* in income

Table 7.1: Overview of analysis sample

<i>Sample</i>	East German couples		West German couples	
	Observations	Individuals	Observations	Individuals
	(1)	(2)	(3)	(4)
All	77,124	8,044	221,887	24,199
Post 1990	77,124	8,044	181,729	20,732
Event time -3 → +6 years to first birth	8,806	1,615	26,743	4,862
Event time & pre- and post-birth observation	5,133	664	18,042	2,508

Notes: Table shows number of observations for different samples and number individuals in the samples. East and West Germans are defined by their 1989 location. All samples cover survey years 1990-2019 where information for both East and West Germans are available. Source: SOEP v36

and a binary indicator for the couple following a main male breadwinner model ($< \frac{1}{3}$ of female income share) are presented in the Appendix. To capture not only the income distribution, but also the degree of participation in the labour market—an aspect of women’s empowerment in itself—I also show results for the female share of weekly working hours in couples.

On the domestic level, I look at contributions to domestic work (child care and housework¹⁶) in the household. Specifically, the questionnaire asks how many hours respondents spend on those tasks on average weekdays.¹⁷ In Appendix section 7.A.4, I compare this time use information with time-use diary data (see next subsection) to validate the usage of this information in SOEP. Focusing not only on housework but also on child care is particularly important in this context, as child care obligations are often an obstacle to both parents being (full-time) employed. The couple-perspective is a particular advantage for those outcomes, as due to a strongly differing supply of day care in East and West Germany,¹⁸ parents in East Germany have fewer hours of potential child care obligations, unless lower day care is fully compensated by informal care arrangements.¹⁹ Shares of child care within a couple take this into account. Following Siminski and Yetsenga (2020), I also use a proposed household specialisation

¹⁶The questionnaire specifies that housework refers to “washing, cooking, cleaning”. These tasks are commonly defined as *routine housework* as these have to be conducted regularly and are more difficult to postpone (Borra et al., 2021).

¹⁷For both housework and child care I set observations to missing if more than 20 hours per day are indicated. These are 0.02% of observations for housework and 4.6% of observations for child care (both refer to post-birth observations. In 97% of cases when child care is recoded, 24 hours of child care per day are indicated. Perhaps a comprehensible answer, but not suited to this analysis.)

¹⁸As of March 2020, 52.7% of under threes were enrolled in East Germany compared to 31% in West Germany. See: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Soziales/Kindertagesbetreuung/Tabellen/betreuungsquote-2018.html>, last accessed July 21, 2021.

¹⁹Looking at children aged one to six, I find that West German parents spend on average 1.4 hours more on child care per weekday.

index (SI_2 in their paper) to summarise the division of market and domestic work within the household in one number.

$$SI = \frac{DW_F}{DW_F + DW_M} - \frac{MW_F}{MW_F + MW_M} \quad (7.1)$$

DW and MW denote domestic and market work, respectively, and the subscripts indicate female and male contributions per unit. The index ranges from -1 (non-traditional specialisation) to 1 (traditional specialisation, i.e. the woman is solely responsible for domestic work and the man for market work²⁰) with 0 implying equal contributions to both domains by the partners. The distribution of SI in East and West German couples is presented in Appendix Figure 7.A.2.

Table 7.2 displays pre-birth characteristics of the sample. West Germans in the sample are about two years older, more likely to be married and have substantially higher pre-birth earnings.

7.3.2 Additional sources

Time-use data I additionally use two time-use surveys from Germany. The first one is a time-use survey from the GDR conducted in 1985 and 1990 (before reunification) by the statistical office of the GDR. Tasks were recorded over 24 hours on a pre-determined day. Participating households were also part of a representative household finances study and the data is representative for worker and employee households (Fiebiger, 1991). Reliable micro data from the GDR is rare, so this data source offers a unique opportunity to gain insights on time use and gender inequality in a state-socialist country where participation and working hours are relatively fixed. Berkes et al. (2021) provide further details on the data.

Second, I use three waves from the (post-reunification) German Time-Use Survey taken in 1991/92, 2001/02 and 2013/13. All adult household members record three-digit classified activities in ten (five in 1991/92) minutes slots over three (two in 1991/92) diary days (Maier, 2014). In Appendix section 7.A.4 I compare average values obtained in the time-use survey and from SOEP using the same survey years and comparable samples (Borra et al., 2021, do a similar validation of time-use and survey data for the UK and US). In the time-use survey I can distinguish between households' current location in East and West Germany, but no information is given on the place of birth or socialisation of individuals.²¹

²⁰Hereby I follow Farré and Vella (2013) in using the term *traditional* when referring to a situation when women are responsible for domestic work and men for market work.

²¹Appendix Table 7.A.5 shows that estimates from the SOEP based on socialisation or current location are indistinguishable.

Table 7.2: Pre-birth characteristics

	East German couples		West German couples	
	Women (1)	Men (2)	Women (3)	Men (4)
<i>Individual characteristics</i>				
Age in years	26.56 (4.03)	29.75 (5.03)	28.97 (4.42)	32.00 (4.97)
Current location in East Germany	0.83 (0.38)	0.80 (0.40)	0.01 (0.11)	0.01 (0.11)
Married	0.41 (0.49)	0.35 (0.48)	0.62 (0.49)	0.61 (0.49)
Higher schooling degree	0.35 (0.48)	0.28 (0.45)	0.40 (0.49)	0.44 (0.50)
University degree	0.24 (0.43)	0.18 (0.39)	0.23 (0.42)	0.31 (0.46)
Any employment	0.81 (0.39)	0.89 (0.31)	0.90 (0.30)	0.95 (0.23)
Full-time employment	0.70 (0.46)	0.84 (0.36)	0.76 (0.42)	0.90 (0.30)
Weekly working hours	33.32 (17.85)	39.78 (16.50)	34.92 (14.73)	41.45 (13.03)
Monthly gross earnings	1,413.92 (1026.78)	1,989.96 (1268.72)	2,038.29 (1194.80)	3,101.88 (2204.66)
Hourly wage	10.75 (4.97)	13.15 (5.75)	14.57 (5.98)	18.93 (9.81)
Daily hours of housework	1.73 (1.41)	0.81 (0.77)	1.70 (1.25)	0.82 (0.71)
<i>Couple characteristics</i>				
Female share of labour income	0.41 (0.26)		0.41 (0.22)	
Female share of working hours	0.44 (0.25)		0.46 (0.21)	
Specialisation index	0.24 (0.42)		0.22 (0.37)	
Observations	762	709	2389	2247

Notes: Table shows pre-birth (1 to 3 years) characteristics separately for women and men of East and West German couples (by their 1989 location). Higher schooling degree denotes university entrance qualification (*Abitur*). Earnings and wages reported in 2010 Euros. Specialisation index defined as in equation 7.1. Source: SOEP v36

In both time-use surveys the analysis focuses on different-sex couples of working age. As both data sets are cross-sectional, no information on future fertility can be used. Thus to approximate the impact of children, I use childless couples of a similar age range as a comparison group (see next subsection).

pairfam The analysis of attitudes is conducted with data from the German Family Panel *pairfam*. The longitudinal household survey with a focus on researching partnerships and family dynamics has been conducted annually since 2008 with 11 waves

released to date.²² Similar to SOEP, the same set of respondents are interviewed in every annual survey wave, due to which birth events are often observed in the data. Respondents are asked about a wide range of attitudes in every survey year, thus allowing to implement event study estimates to analyse whether the arrival of children is associated with a change in attitudes of individuals.

7.3.3 Empirical approach

To analyse the dynamic effect of having children I employ an event study specification following Kleven et al. (2019b):

$$y_{ist}^r = \sum_{j \neq -1} \alpha_j^r \cdot \mathbb{I}[j = t] + \sum_k \beta_k^r \cdot \mathbb{I}[k = age_{is}] + \sum_y \gamma_y^r \cdot [y = s] + \epsilon_{ist}^r \quad (7.2)$$

for outcome y of individual (couple) i , of region $r \in \{East, West\}$, in year s , and event time t . Standard errors are clustered at the individual level. Event time $t = 0$ denotes the 12 months after a couple's first child is born. The event time coefficients $\hat{\alpha}_t^r$ are normalised to the pre-birth year and indicate how the outcome variable dynamically evolves relative to the counterfactual of not having a (first) child. By including age and survey year dummies, the $\hat{\beta}$ s and $\hat{\gamma}$ s non-parametrically net out life cycle trends and time trends such as concave age-earnings profiles due to return to experience or economic shocks in certain years.²³ Identification stems from variation in age at first birth and across time. Equation (7.2) is estimated separately for East and West German couples, to allow for differential life cycle or time effects.²⁴ An attractive feature of event study designs is that obtained coefficients can be presented neatly in event study graphs. As the main estimates do not condition on future fertility, coefficients capture the total effect of children on gender inequality and differences between East and West German parents.

Besides the event study specification, similar to Kuziemko et al. (2018) I also estimate a simpler difference-in-differences-type equation to obtain a summary coefficient for the average post-birth effect. I specify three discrete points in time; pre-birth, the year of

²²A documentation of the latest release is provided by Brüderl et al. (2020) and a detailed description of the study is found in Huinink et al. (2011).

²³If life cycle and time effects were not taken into account, the event-coefficients would simply correspond to mean values for the event time relative to the pre-birth year as in Appendix Figure 7.A.3. The figure reveals a small drop in female working hours in the year before birth, which may be due to anticipated fertility or if mothers have entered maternity leave already.

²⁴Differences could for example arise if due to different socialisation if it is the norm in one region to have children only after a few years of labour market experience.

birth and post-birth. The equation is

$$y_{ist}^r = \zeta^r \cdot birth + \delta^r \cdot post + \sum_k \phi^r \cdot \mathbb{I}[k = age_{is}] + \sum_y \theta_y^r \cdot [y = s] + u_{ist}^r \quad (7.3)$$

The coefficient of interest, δ , is reported in all event study graphs as well. While the event study estimates based on equation (7.2) can tease out the detailed evolution of effects by year, an advantage of this estimation technique is that it provides one summarising coefficient and due to pooling of several years it requires fewer annual (event time) observations, making it more suitable for looking at subgroups. All descriptive analyses in this paper are based on calculations using survey weights.

The impact of children on a wide range of outcomes can most credibly be estimated in an event study framework and this has become the standard in the literature. However, in some cases due to data limitations this is not possible; event study estimates crucially rely upon a panel structure to be able to control for pre-birth realisations of the outcome variables.

To be able to assess time use in more detail, I additionally use time-use data from the GDR and the German Time-Use Survey, which are both repeated cross-sections. In contrast to simply documenting East-West differences as has been extensively done in the literature, the aspect of interest here is whether the arrival of children exacerbates such differences. The sociological literature has argued that parenthood can activate gender norms (sleeping effect) and lead to more gender-traditional attitudes (e.g. Corrigan and Konrad, 2007; Cunningham, 2001).

To approximate the impact of children in cross-sectional data, I compare outcomes of couples with young children to childless couples of a similar age range. Specifically, I use a sample of couples aged in the 5th to 95th percentile of first-time parents. I estimate the equation

$$y_{ist}^r = \kappa \cdot child + \sum_k \omega^r \cdot \mathbb{I}[k = age_{is}] + \sum_y \lambda^r \cdot [y = s] + X' \omega + \nu_{ist}^r \quad (7.4)$$

where *child* is a binary indicator equal to one for couples having a child aged one to six years, and zero for childless couples. X contains indicators for higher education and marital status. In cross-sectional data one cannot assess the validity of this control group, as of course only some of these couples will become parents and the problem of selection into parenthood arises. I use SOEP (panel-)data to check how well this approach fares compared to event study estimates. Appendix Table 7.A.2 shows post-birth coefficients based on equation (7.3) and contrasts them with those obtained from equation (7.4) with the approximated control group. For the outcomes female share

of income, working hours and housework, true and approximated coefficients are very close, only for female share of child care is the difference in East German households notable. Despite those encouraging results, due to the imperfect control group those results ought to be taken with a pinch of salt; rather than showing the impact of children for couples with children (an average treatment effect on the treated), these are conditional differences between couples with and without children of similar age.

7.4 Results

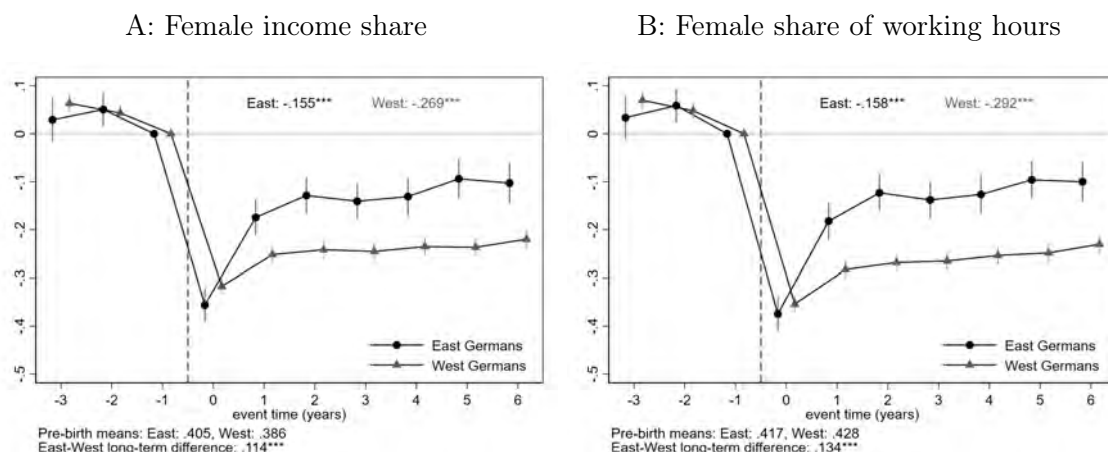
7.4.1 Labour market and domestic outcomes - event study

Labour market outcomes Figure 7.3 shows the impact of children for the two main couple-level labour market outcomes across event time by region. Coefficients are normalised to the pre-birth year ($t = -1$), range lines indicate 95% confidence intervals calculated with standard errors clustered at the individual level. Panel A shows the impact on female income share. In the year after birth ($t = 0$) the shock to the female income share is similar in East and West German couples. Afterwards the share in East German couples recovers strongly, but almost stagnates in West German couples at 26.9 pp averaged over the post-birth years. The impact in West German couples is 74% larger relative to East German couples, where the share is reduced by 15.5 pp. As can be seen in the pre-birth averages, this is by no means a move to the same post-birth value in East and West Germany (say 25%) but a further divergence in the earnings share between East and West German couples.

A potential explanation for these differing child penalties could be that bargaining power—through earnings potential—of women in West German couples is notably weaker. However, when restricting the analysis to couples with *higher* female pre-birth earnings, where additionally only 13.5% of women have lower educational attainment, the East-West long-term difference still amounts to 9.6 pp. Additionally, differential future fertility could exacerbate differences if more West German couples have additional children. Yet looking only at one-child families, East-West differences amount to 11.9 pp, refuting this mechanism. Results by number of children and the local effect of having a second child are shown in Appendix Figure 7.A.4. For one child-families, the female income share in East German couples recovers after four years, but remains more than 10 pp lower for West Germans. The local effect of a second child is similar for East and West Germans.

In Panel B of Figure 7.3 the share of working hours is displayed. The similarity of the impact on those two outcomes indicates that the effect on hourly wages differs little. In fact, the negative impact on (log) hourly wages is slightly more pronounced

Figure 7.3: Impact of children on labour market gender inequality



Notes: Figure shows event study estimates for the respective outcomes. Units of observations are couples. Coefficients are normalised to the pre-birth year ($t = -1$), means from this year are displayed in the figure notes. Long-term coefficients shown in the figure stem from estimates pooled over post-birth years ($t = 1-6$). Income share refers to gross monthly labour income. East and West Germans are defined by their 1989 location. Significance levels: * < 0.1 ** < 0.05 *** < 0.01 . Source: SOEP v36

for East German mothers where there is less selection into post-birth employment than for West German mothers. Appendix Table 7.A.3 shows overall couple-level sums for the main outcomes, from which the shares are calculated.

While having a child is a permanent negative shock to gender equality in the labour market for couples from both regions, it is so to a much larger degree in West German couples. Additional results are presented in the Appendix. To include couples were both have zero earnings or hours (3-4% of observations) instead of *shares* I also show *gaps* for the outcomes (Appendix Figure 7.A.5). As due to different labour market conditions the earnings level between East and West Germany differs strongly, for ease of interpretation shares are preferred over gaps. However, results are qualitatively the same. Panel C in Appendix Figure 7.A.5 further shows the effect of a discrete binary categorisation of the couple having a main male breadwinner ($< \frac{1}{3}$ of female income share), which increases in West Germany by 58.3 pp compared to 32.3 pp in East Germany. Individual-level event study estimates are presented in Appendix Figure 7.A.6. In line with existing evidence, fathers' labour market trajectories are not strongly affected by children, with small negative coefficients observed for East German fathers.

Domestic work Next, I turn my attention to non-labour market outcomes and look at contributions to domestic work, i.e. housework and child care. Although East German mothers' weekly working hours recover to some degree from two years

after birth onward, it does not follow unambiguously that their relative contributions to domestic work decrease accordingly. If one partner, mostly mothers, temporarily withdraws from the labour market, additional domestic work and especially child care is often covered by this person. Strongly unequal division of such work may lead to lock-in effects if the partner with the longer absence continues to be the main caregiver even after returning to the labour market, which could have negative long-run effects on working hours and flexibility, which negatively affects wages (Goldin, 2014).

Estimates for domestic work are obtained using the same framework as the previous section, but for child care the estimation cannot follow the identical event study logic as child care investments only start when the child is born.²⁵ To be consistent the results for child care are still shown in the same way, but the normalisation to $t = -1$ is irrelevant. To a lesser, non-deterministic degree this also holds for housework because the inputs required post-birth increase strongly and this holds even more when more time is spent at home. For workings hours, in contrast, both pre- and post-birth the choice set is in the same fixed range, say 0-50 weekly hours. I also present estimates for the specialisation index by Siminski and Yetsenga (2020) described in subsection 7.3.1, which indicates to what degree couples divide market and domestic work on a continuous scale from a *non-traditional* ($SI = -1$) to a *gender-traditional* specialisation ($SI = 1$). The advantage of such an index is that it summarises distinct aspects of household specialisation in one number. Because it is calculated with shares, it is less prone to distortions due to overall level differences between regions (e.g. labour market conditions or day care availability).

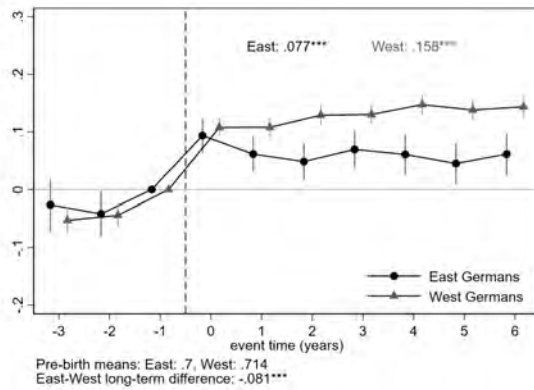
Figure 7.4 presents the results for domestic work. Two aspects stand out. First, the pre-birth means for housework (Panel A) indicate that, in contrast to earnings and workings hours where prior to children the distribution was more equal, gender inequality in this domestic domain was already prevalent without children as women were on average responsible for around 70% of housework. Post-birth, when the total amount of housework increases as well in couples (Appendix Figure 7.A.3 and Appendix Table 7.A.3), the female share increases by 8-16 pp with—as for labour market outcomes—a stronger effect in West German couples.

Child care (Panel B), starting from a base of zero, is (still) a predominantly a female domain with an initial share of 80% in East and West German couples and only a decrease of just below 10 pp in East German couples as the child gets older. The

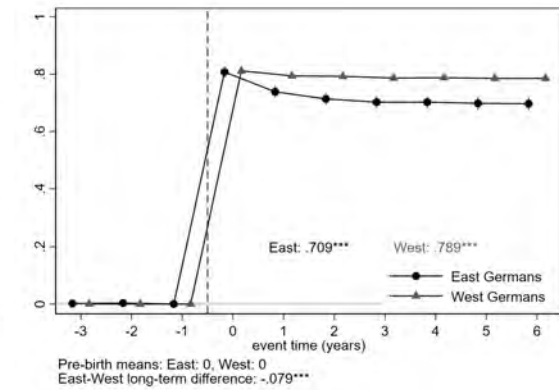
²⁵ The questionnaire asks for child care in general and not necessarily for child care of the respondent's children. Yet I observe that pre-birth the average daily time spent on child care is less than 8 minutes for women in the pre-birth year (2% of women report time spend on childcare) compared to 9 hours in first post-birth year. Due to this I am confident that child care time measures to large degree time with the own child and I set pre-birth child care time to zero.

Figure 7.4: Impact of children on domestic gender inequality

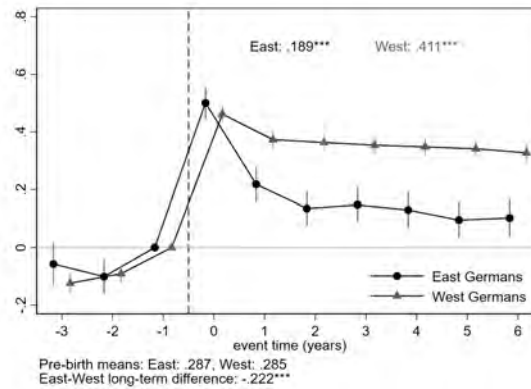
A: Housework share



B: Child care share



C: Specialisation index



Notes: Figure shows event study estimates for the respective outcomes normalised to the pre-birth year ($t = -1$). Housework and child care refer to shares on weekdays. The specialisation index is defined in equation (7.1). See Figure 7.3 for other notes. Source: SOEP v36

constant high share in West German couples is even more remarkable as the total amount of daily child care in couples decreases from almost 11 hours ($t = 1$) to less than 9 hours ($t = 6$), meaning that the decreased total time is decreased proportionally by fathers as well who from the onset had much lower involvement. The relative contributions to market and domestic work are summarised in the specialisation index in Panel C. While couples were already specialised pre-birth (0.287 in East and 0.285 in West German couples), the arrival of the first child leads to a spike in specialisation in couples with relative increases in the index of 66 and 144%. Remarkably, even 15 years later the impact of the first child in West German couples on the specialisation index is 0.30 (0.12 in East German couples, not shown but available upon request). This

suggests that having a child leads to a *permanent* traditional orientation in couples.²⁶ Estimates for *total hours of work* per weekday (paid work, housework and child care) in Appendix Figure 7.A.8 show that the female share of overall work increases, and more so for West Germans, meaning that the increase in domestic work is not fully offset by a decrease in paid work.

As for the labour market outcomes, event study estimates in gaps (Appendix Figure 7.A.9) and the individual-level contributions (Appendix Figure 7.A.10) are displayed in the Appendix. Because the overall levels of housework and child care change strongly by event time, an aspect that is less visible when focusing on shares, is that absolute gaps show even stronger divergences within couples after the arrival of a child. Additionally, East-West differences are also stronger with a continuously increasing housework gap in West couples. For child care, the differences in the impact on the gap 6 years after having a child is about three hours per day. At this age, in both West and East Germany almost all children attend day care or school.

Long-run estimates, i.e. average estimates for one to six years after birth, and standard errors of estimates for the main labour market and domestic outcomes are summarised in Table 7.3. Columns (5) and (6) show the difference in long-run estimates between East and West German couples obtained from a fully interacted model. The table also shows estimates from regressions with additional pre-determined characteristics (see table notes) in even-numbered columns to control for potentially confounding factors. Coefficients are generally stable when control variables are added in the estimation and support strong East-West differences in the long-term effects of children on within-couple gender inequality.

7.4.2 Time-use evidence

A downside to the usage of survey data for analysing time use is the inherent lack of precision (SOEP only allows for answers in full *hours*), recall bias, the issue of social desirability, and measurement error. Data from time-use studies, recorded in fine-grained diaries over survey days, resolve those issues and are generally considered to be more accurate, especially for activities other than paid work that are conducted in less regular intervals (Kitterød et al., 2005). Time-use researchers have found that despite differences in activities in diary versus survey data, the approaches tend to yield

²⁶Estimates in Figure 7.4 are based on information for weekdays, when the trade-off between market work and domestic work is higher. The survey also asks for contributions on both weekend days biannually. Estimates using the weekly information are presented in Appendix Figure 7.A.7. Results indicate that child induced inequality in domestic work is lower on weekend days as the coefficients are slightly reduced, but they remain large and East-West differences are very close to the main estimates.

Table 7.3: Long-run impacts of children

	East German couples		West German couples		East-West difference	
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Female income share</u>						
Long-term effect	-0.155*** (0.016)	-0.140*** (0.016)	-0.269*** (0.007)	-0.241*** (0.008)	0.114*** (0.017)	0.102*** (0.018)
<u>Female share of working hours</u>						
Long-term effect	-0.158*** (0.015)	-0.150*** (0.015)	-0.292*** (0.007)	-0.269*** (0.008)	0.134*** (0.016)	0.119*** (0.017)
<u>Female housework share</u>						
Long-term effect	0.077*** (0.013)	0.060*** (0.013)	0.158*** (0.007)	0.129*** (0.007)	-0.081*** (0.015)	-0.069*** (0.015)
<u>Female share of child care</u>						
Long-term effect	0.709*** (0.008)	0.704*** (0.008)	0.789*** (0.003)	0.783*** (0.004)	-0.079*** (0.009)	-0.079*** (0.009)
<u>Specialisation index</u>						
Long-term effect	0.189*** (0.024)	0.167*** (0.024)	0.411*** (0.012)	0.365*** (0.013)	-0.222*** (0.026)	-0.198*** (0.027)
Age, survey year FEs	Y	Y	Y	Y	Y	Y
Additional controls		Y		Y		Y
Observations	4,088	4,026	12,552	12,163	16,640	16,290

Notes: Table shows long-run coefficients ($t = 1 - 6$) of the arrival of children on within-couple gender inequality. Columns (1), (3) and (5) are estimates shown in Figures 7.3 and 7.4. Additional control variables added in other columns: schooling and university degree, federal state dummies (16), migrant background, municipality size class dummies (7) and an indicator for married couples. Standard errors clustered at the couple-level in parentheses. Significance levels: * < 0.1 ** < 0.05 *** < 0.01. Source: SOEP v36

comparable patterns between groups and are therefore insightful (see, e.g., Baxter and Bittman, 1995; Marini and Shelton, 1993). Due to the lack of a panel structure in German time-use studies, the impact of children cannot be estimated with this data, but it allows for a more detailed inspection of gender inequality in time use in couples with and without young children. These analyses are complementary to the event study estimates using SOEP in order to gain a thorough understanding of within-couple gender inequality, particularly for non-market work, and to include evidence from the GDR.

Since reunification, three time-use surveys (1991/92, 2001/02, 2012/13) have been conducted in Germany. Additionally, two time-use surveys from the last years of the GDR (1985 and 1990) offer a unique opportunity to study gender inequality in a social-

ist system, where, generally speaking, individuals were obliged to work and differences in working hours between men and women were much smaller. Despite some differences in the sampling design, the time-use surveys have been conducted in a comparable fashion in the GDR and in reunified Germany. This allows to compare outcomes from the GDR and to those from East and West Germany in a consistent way, which was not feasible with the SOEP. To contrast couples with and without children, the sample is restricted to couples with either i) children under 6 or ii) couples with no children in the household but a female age distribution in the range of the 5th to 95th age percentile of those with children (see subsection 7.3.3).

In Figure 7.5 the couple-level distributions of the female share of (paid) working hours, housework and child care, and the specialisation index are plotted separately for the GDR, East and West Germany. In Panel A the narrow distribution of working hours in the GDR is apparent;²⁷ of couples with both partners working, the female share lies in the range of 0.4 and 0.6 in 74% of couples. Additionally, with 5 pp the difference between couples with and without children is quite low. After reunification in Germany, the difference in the working hours distribution by children in East Germany is larger than in the GDR, but much smaller than in West Germany (14 vs. 22 pp). Overall a wider distribution of the working hours share is evident in reunified Germany, which is mostly due to a larger share of individuals not in employment.

With market work being relatively evenly distributed in the GDR (and to a slightly lower degree later in East Germany), strong gender inequalities can be observed in domestic work as documented by Nickel (1992). About two-thirds of housework in the GDR is performed by women, but the average differs little by children. The housework shares for childless couples after reunification are quite similar in both regions of Germany, but in line with the stronger decrease in working hours, children increase the female housework share substantially and more so in West Germany.

Child care is mostly the responsibility of mothers across space and time. Notably, the share in the GDR is more than 10 pp *higher* in the GDR than in East and West Germany. Gender egalitarian policies in the GDR focused on labour market aspects (Cooke, 2007), but in terms of domestic work—and especially child care—the data does not suggest that this had any spillovers on an overall more gender egalitarian distribution (Berkes et al., 2021).

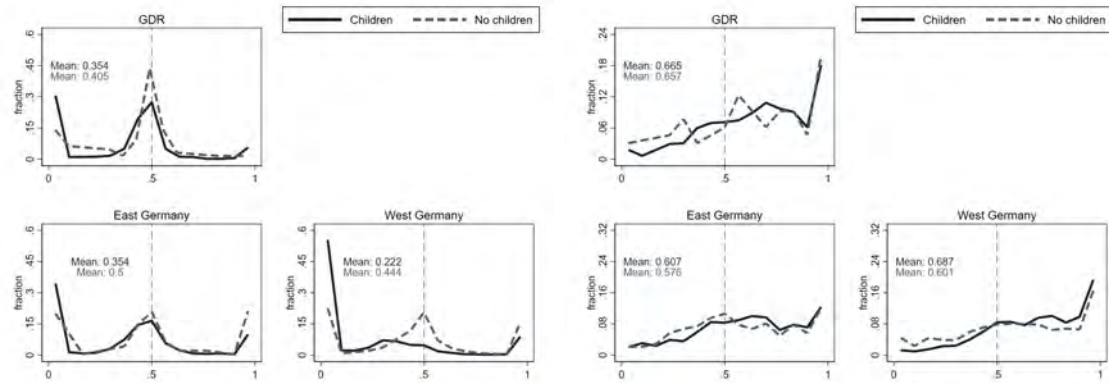
The specialisation index summarises the gender-specific specialisation in households and illustrates that the GDR was in fact more gender egalitarian than West Germany, but less so than post-reunification East Germany. However, the difference by children

²⁷By law, a standard work week was 43.75 hours, and 40 hours for mothers with two children below the age of 16 (Rosenfeld et al., 2004).

Figure 7.5: Distribution of activities - time-use data

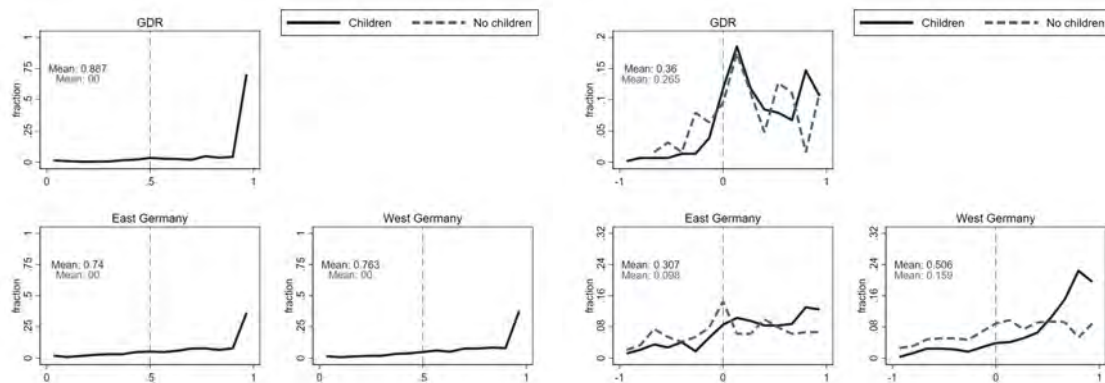
A: Share of working hours

B: Share of housework



C: Share of child care

D: Specialisation index



Note: Figures show within couple shares of respective activities per survey day. Sample is restricted to weekdays. Region refers to current location of couples. GDR data from 1985 and 1990, East and West German data from 1991/1992, 2001/02 and 2012/13. Children indicates a child under 6 years in the household, couples with no children are in the 5th-95th percentile age range of couples with children in the sample. Distribution calculated in 15 bins of equal width. Sources: Time-Use Study of the GDR and German Time-Use Survey

in the GDR is smaller. In Appendix Table 7.A.4 conditional differences controlling for survey wave and life-cycle effects are presented. To make the results more comparable to the survey results using the SOEP, the table additionally includes coefficients for routine housework. The “impact” of children on market work and housework is generally smaller using time-use data, which could be attributed to different samples, measurement and the imperfect comparison of households with and without children. Regardless, East-West differences remain strong.

7.4.3 Attitudes

Differences in gender-related attitudes between East and West Germans have been well documented in the literature (e.g. Bauernschuster and Rainer, 2012; Zoch, 2021)

with East Germans persistently holding more egalitarian views. Building on this, this subsection examines how attitudes specifically related to maternal employment differ and whether the arrival of children has an impact on such attitudes. This section uses data from the German family panel *pairfam*. A set of questions ask parents to what extent mothers of children of different age groups should ideally be working. An attractive feature of this is that it allows to analyse differences in attitudes towards both the extensive and intensive margin of maternal employment by child age.

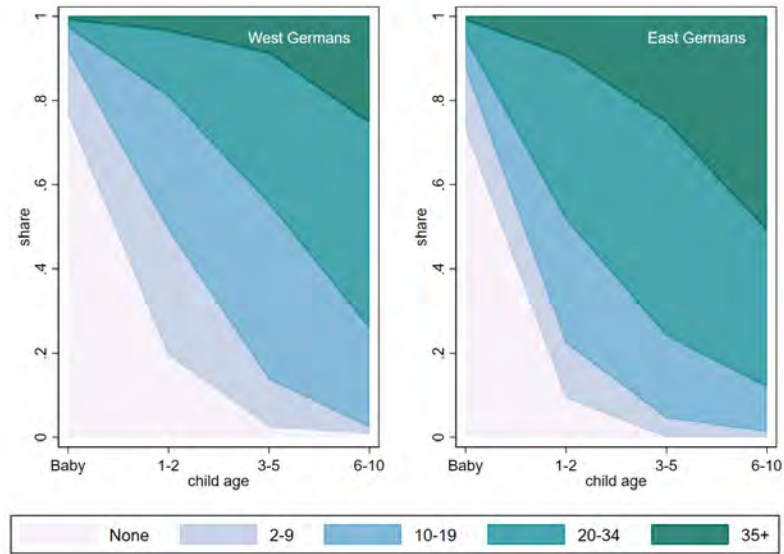
The distribution of ideal working hours by child age is presented in Figure 7.6, Panel A. In the first year of a child, both East and West Germans indicate that mothers should not be in employment or, if so, only be working few hours. This is consistent with very similar effects on labour market outcomes in the first year post-birth (Figure 7.3). Then, however, attitudes towards maternal employment begin to diverge; a smaller share of West Germans indicate that mothers should not be working at all, but most respondents are only in favour of part-time work with moderate hours. In contrast, about half of East Germans suggest that mothers of children aged 1-2 should be working 20 hours or more per week (which only 19 percent of West Germans are in favour of). With increasing child age, longer maternal working hours are deemed ideal among East and West Germans, but even for more gender-egalitarian East Germans less than half of respondents prefer full-time maternal working hours. If those attitudes are adhered to, a full catch-up of mothers in terms of labour market outcomes is incompatible. Panel B of Figure 7.6 displays East-West differences in attitudes towards working hours by child age, making apparent that differences are initially small and with increasing child age are first stronger at the extensive and later at the intensive margin.²⁸

While this evidence is intriguing as it helps to explain differential recovery for East and West Germans, it is unclear whether these differences were pre-existent and constant, or either exacerbated or diminished after the arrival of children. Kuziemko et al. (2018) have documented that mothers in the US underestimate the effect of having children on their future labour supply, a finding they denote as "the mommy effect". In line with this, attitudes towards maternal employment may change after the arrival of children. A partial convergence (or further divergence) of East-West attitudes could occur if upon becoming parents, East Germans find parenthood harder than expected and culturally induced favourable attitudes towards maternal employment are reduced (or vice versa). To investigate this, I focus on two questions which are asked in the survey irrespective of respondents having children, allowing to investigate whether these attitudes change. *Women should be more concerned about family than about career*

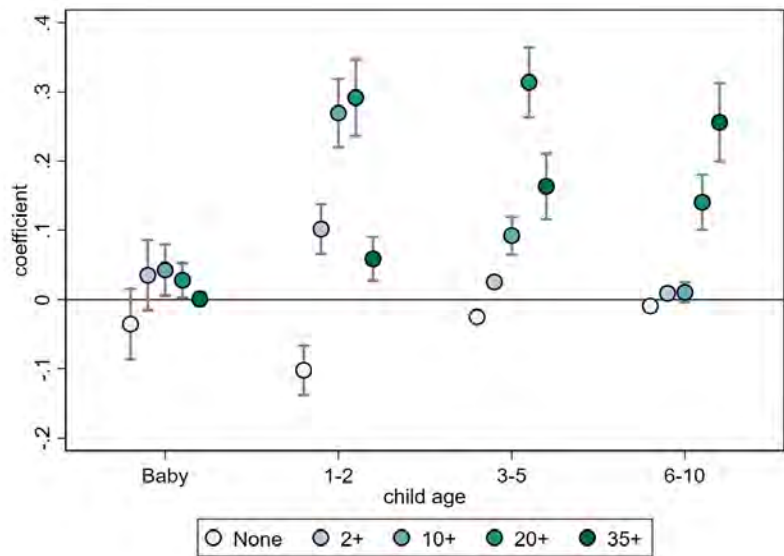
²⁸This pattern can also be observed for *desired working hours of respondents* using SOEP data (Appendix Figure 7.A.11).

Figure 7.6: Maternal employment by child age

A: Ideal weekly working hours for mothers



B: East-West differences



Note: Panel A shows the distribution of indicated ideal working hours for mothers of children of different ages. West and East Germans are assigned according to their country of birth (GDR or FRG). Panel B shows coefficients and 95% CIs of East-West differences. The underlying questions are only asked to respondents with children. Source: pairfam waves 1-11

and *A child under age 6 will suffer from having a working mother*. Both variables are coded from 1 (disagree completely) to 5 (agree completely), for ease of interpretation

both variables are used as binary indicators if respondents indicate partial (4) or full (5) agreement.

Table 7.4: East-West differences in attitudes and the impact of children

	Women should be more concerned about family than career (0/1)			Child under 6 will suffer with working mother (0/1)		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A: East-West differences</i>						
Mean of dep. variable	0.156	0.158	0.153	0.213	0.161	0.268
East dummy	-0.058*** (0.010)	-0.059*** (0.014)	-0.060*** (0.014)	-0.143*** (0.011)	-0.106*** (0.013)	-0.182*** (0.017)
Sample	Pooled	Women	Men	Pooled	Women	Men
Wave & age FEs	Y	Y	Y	Y	Y	Y
Observations	13,621	7,084	6,536	13,561	7,061	6,499
<i>Panel B: Impact of children on attitudes</i>						
Mean of dep. variable	0.180	0.180		0.184	0.184	
Long-term effect of children	0.062*** (0.019)	0.053*** (0.020)		0.032* (0.019)	0.023 (0.021)	
East dummy	-0.069*** (0.016)	-0.102*** (0.028)		-0.097*** (0.017)	-0.130*** (0.029)	
East x post-birth		0.043 (0.032)			0.043 (0.032)	
Sample	Pooled	Pooled		Pooled	Pooled	
Age & wave FEs	Y	Y		Y	Y	
Observations	3,689	3,689		3,679	3,679	

Note: Panel A presents East-West differences in agreement to statements listed at the top of the table, both coded as binary indicators. All regressions include age and survey wave FEs. Panel B shows the impact of children on those attitudes, and differential effects for East German couples. For those estimates the sample is restricted to three years pre- to six years post-birth of the first child. Long-term effect refers to the average post-birth effect. Standard errors clustered at the individual-level in parentheses. Significance levels: * < 0.1 ** < 0.05 *** < 0.01. Source: pairfam waves 1-11

In an intermediate step, Panel A of Table 7.4 displays East-West differences. In line with the extant literature, East Germans are about 6 pp less likely to agree that women should be more concerned about family than career (37% relative to the sample mean) and 14 pp less likely to say that a child suffers under a working mother (67%). Regarding women putting family over career, gender differences in responses are small (columns 2 and 3). However, men are much more likely to agree with the statement that young children suffer with a working mother (columns 5 and 6), but the larger East dummy for the men-only sample indicates that gender differences in East Germany are generally smaller in this regard.

In a second step, I take advantage of the panel structure and use an event study design as in the main analysis in subsection 7.4.1. Similarly, I use data from three

years pre- to six years post-birth and report the average pooled post-birth coefficient. Panel B, column 1 shows that children are associated with a higher share agreeing that women should be more concerned about their family and that in this sub-sample of (becoming) parents East Germans are also less likely to agree with this. In column 2, the East German dummy is interacted with the post-birth period to elicit whether children lead to a differential effect regarding this attitude for East Germans. While the coefficient is positive, it is not statistically significant at conventional levels. For the question on whether children under 6 suffer with a working mother—working mothers were the norm in the GDR—the evidence for children impacting those attitudes is weaker. However, both the coefficient for children as well as the interaction with East Germans are positive, suggestive of an increase for East Germans (jointly the coefficients are significant at the 5% level).

The examination of attitudes with respect to children and maternal employment in East and West Germany overall supports the notion that attitudes are more egalitarian in East Germany. The arrival of children is associated with more gender-conservative attitudes, with weak evidence for East-West differences becoming smaller (but attitudes certainly remain more egalitarian for East Germans). However, remaining differences in gender- and children related attitudes continue to manifest themselves in child-induced labour market penalties that negatively affect relative labour market outcomes of mothers more strongly in West Germany.

Additionally, I also find that East German couples are more likely to dissolve following child birth (Appendix Figure 7.A.12). In the GDR, single mothers were actively supported, e.g. employers were mandated to preferentially hire them and they were eligible for longer paid parental leave. Universal day care provision with long hours also enabled single mothers to work full-time. In contrast in the FGR, the state provided much less support for single mothers. Differential partnership stability of young parents appears to be yet another legacy of the German division.

7.5 Robustness

In this section, I run a battery of robustness checks to support the hypothesis that the results are primarily driven by socialisation of couples. For this I will use SOEP, as the main analyses rely on this data set, and because its panel structure and richness in variables makes it most suitable to assess robustness. Stability of estimates for the main outcomes examined in subsection 7.4.1 will be shown which then also gives support to the validity of other estimates.

A main concern may be that results are not driven by the couple's origin, but by the current location where households reside.²⁹ I.e. current local norms (or institutions) are more relevant than norms individuals are exposed to during childhood or adolescence. This may then also imply that horizontal cultural transmission through peers is more important than vertical transmission through generations (Bisin and Verdier, 2001). The first two rows of Appendix Table 7.A.5 display outcomes by current location in East and West Germany and results are almost indistinguishable to those presented in Table 7.3. The following rows show DD coefficients for *origin* \times *location* cross combinations. Results for East and West German couples living in their region of origin are again very similar. For East Germans living in West Germany, i.e. couples who have moved, the effect on labour market outcomes are similar to East German stayers (similar results on the individual level have been found by Boelmann et al., 2021; Collischon et al., 2020).³⁰

Another prime candidate to be the driver of East-West differences is the different population share with a migrant background. This may be a relevant factor if individuals born abroad have different gender norms to the native-born population and thus respond differently to the arrival of children. In 13.6% of West German couples, but only 2% of East German couples are both partners born abroad. Results with the sample restriction of both partners born in Germany are displayed in Appendix Table 7.A.6. Coefficients are statistically identical to the main results.

In the current analysis, mixed couples, i.e. those where either partner lived in the GDR and the other in the FRG in 1989, are excluded from the analysis. These are 6.6% of all couple-level observations. Like Lippmann et al. (2020), I find that descriptively these couples are between pure East and West German couples with lower child penalties when the man is from East Germany. If these mixed couples are assigned to either East or West German couples, this has only minor influence on the estimates.

Appendix Figure 7.A.13 subsequently excludes each of the 16 federal states to corroborate that effects are not driven by a specific state. Note that as this exercise is based on the current location of couples, I use these estimates as the reference point. These were shown to be very similar to those based on both partners' socialisation (see above). Coefficients in Appendix Figure 7.A.13 indicate stability to the exclusion of states with estimates being slightly larger when the East German state of Saxony is excluded.

Next, I split the ten West German federal states in all possible combinations to belong to either of the two 'treatment groups'. Then I estimate event study estimates

²⁹This would then directly devalidate the analysis based on time-use surveys as these only rely on household's current location.

³⁰Too few West German couples live in East Germany to conduct statistical analysis with this sample.

for those two groups and calculate the difference in post-birth coefficients and contrast these to the actual East-West difference observed. This placebo exercise gives an indication how likely these regionally differential responses to the arrival of children could have arisen if Germany had been divided in another way along state borders. Appendix Figure 7.A.14 shows histograms of the coefficients and the observed East-West difference indicated by a vertical dashed line (based on current location). For the five main outcomes, none of the estimates exceeds the observed East-West difference.

A reason why the child penalty for women is smaller in East Germany could also be that worse economic conditions in East Germany frankly demand both partners to return to employment quicker. In 2018, GDP per capita in East Germany was only 75% of the West German level (43% in 1991). If better economic conditions in some parts allow young families not to have both parents working (full-time), specialisation into market and domestic work may be easier feasible.³¹ To test this hypothesis, I split West German counties into low- and high-income counties (by GDP per capita). To ensure that couples are always assigned to the same group, I use GDP data from 2008 to split counties. Lower income West German counties have only a 5% larger GDP per capita than the average East German county, GDP differences between those counties are thus larger than for the East-West comparison. Event study estimates by GDP are shown in Appendix Figure 7.A.15. No meaningful differences between the groups are observed.

In a similar spirit, day care availability could be a key driver of differences. Day care shortages are prevalent in Germany, especially for under threes (Jessen et al., 2020a), potentially posing a limiting factor for employment. A drawback with analysing this aspect is that differences between East and West Germany are so large—in 2020, the county with the lowest share enrolled in East Germany still exceeded the highest West German county—that West German counties cannot be split to mimic East German counties in this regard. As day care provision for under three was very low in West Germany before the mid-2000s, I only use births after 2004 and (median) split West German counties by day care enrolment. Appendix Figure 7.A.16 shows that overall differences are quite small with long-run effects on the female income share of 21-25 pp, but the evidence is suggestive that higher day care availability is associated with slightly lower penalties. As average differences in enrolment between these counties are only 8%, I compare these numbers with East German births *before* 2006 when enrolment was 30-40%. The long-run penalty on the female income share is 17.1 pp.³²

³¹A traditional specialisation would then make (economic) sense if men have notably higher pre-birth earnings.

³²Looking at opening of child care facilities in Bern (Switzerland), Krapf et al. (2020) find that child care availability does reduce the child penalty for mothers, but with a dampening effect of the

It is worth noting in this context, that results from East Germans who had moved to West Germany, i.e. to a region with lower day care provision, had labour market child penalties much closer to East Germany ‘stayers’ than to their West German peers (Appendix Table 7.A.5). Labour supply elasticities of day care expansion on maternal labour supply are also informative on this matter; research by Müller and Wrohlich (2020) on the effects of day care expansion for toddlers on maternal labour supply in Germany has found elasticities of about 0.2. Under average differences in day care provision for toddlers in the sample period (about 30 pp), this implies that—assuming linearity—only about 40% of the average difference in employment (15 pp) between East and West German mothers of toddlers would be closed by this.³³ Analysing the relationship between the expansion of day care and mothers’ return to work separately for East and West Germany, Zoch and Hondralis (2017) only find evidence for a small effect after the birth of a second child in West Germany. While the importance of day care in explaining part of the East-West differences documented cannot be ruled out, the evidence suggests that this is by far not the sole driving factor of different child penalties.

In a final step, I consider the spatial dimension of East-West differences. Recent research has highlighted pre-existing average differences between the East and West German population before the GDR and FRG were formally established in 1949 (Becker et al., 2020). If these are sufficiently large, differences in modern outcomes may be (predominantly) attributed to those pre-existing differences. Estimating a spatial RD in proximity to the border allows to smoothly control for such gradients, assuming those differences did not jump discontinuously at the later border. Campa and Serafinelli (2019) and Lippmann et al. (2020) follow similar strategies in their analysis of East-West differences.

A rigorous implementation of a spatial RD proves difficult due to the large density of observations required in vicinity of the discontinuity. However, the estimation can follow the intuition of a spatial RD by estimating child penalties in 120km bins around either side of the border.³⁴ Appendix Figure 7.A.17 shows German counties on either

penalty on earnings by 4.5 percentage points (6.3%), the effect is moderate. Kleven et al. (2020) find no evidence that day care expansions in Austria had any effect on child penalties.

³³Bauernschuster and Schlotter (2015) identify an elasticity of 0.37 for children aged three to four years using the introduction of a legal claim and an expansion for this age group in the 1990s.

³⁴The bins are chosen to cover the entirety of East Germany and for each bin to contain at least 1,000 observations (the restriction binds in less densely populated East Germany.) The county furthest away from the border is Spree-Neiße in Brandenburg with a distance of 228km. Campa and Serafinelli (2019) and Lippmann et al. (2020) are able to use finer bins of about 5 and 10km respectively. The reason is that these papers display *average values* based on the entire working age population in their RD plots, whereas the focus here is on data-demanding event study *estimates* for a sample of couples in the years surrounding childbirth.

side of the (former) inner border that are included in the estimation. Coefficients of the effect of children for the five main outcomes are plotted in Appendix Figure 7.A.18. They give no indication that estimates converge in proximity to the border, but in fact larger differences in this region.

7.6 Conclusion

For 41 years Germany was divided into two states with vastly different policies regarding maternal employment. In the GDR, mothers returned to employment quickly, whereas in the FRG policies favoured a (one and a half) male breadwinner model. Since reunification in 1990, East and West Germans are exposed to the same policy environment, but differences in socialisation continue to play a role. This paper examines how child penalties differ between couples who grew up in either in the GDR or FRG but have children in reunified Germany.

The child penalty on the female income share is significantly smaller in East German couples (11.4 pp). Looking at contributions to non-market work, I additionally show that in West German couples the arrival of children is associated with stronger increases in the female share. These findings are in line with more gender traditional attitudes towards maternal employment in the West. The exclusion of numerous potential explanatory factors gives support to the interpretation that differences in norms are a key factor in explaining smaller child penalties on gender inequality in East German couples.

Despite important progress in reducing gender inequality over the past decades across high-income countries, differences in earnings persist and women continue to contribute larger shares to non-market work. Important contributions have found that a large share of remaining gender inequality is child-related (Córtes and Pan, 2020; Kleven et al., 2019b). It is thus of crucial importance to better understand why individuals respond differently to the arrival of children.

The case of East and West Germans couples suggests that norms due to different cultural upbringing play an important role. Besides individually held norms on maternal employment and gender roles more generally, the legacy of GDR policies may also have led society to be more family friendly by reducing the stigma on working mothers. While better day care provision in East Germany was not found to be the main explanatory factor for lower child penalties in East German couples, all day schooling and more family friendly employers may also play a role.³⁵

³⁵Results for East German couples living in West Germany were found to be similar to East German stayers (albeit estimated with a small sample), so the relative importance of these factors is less clear and an interesting aspect for future research.

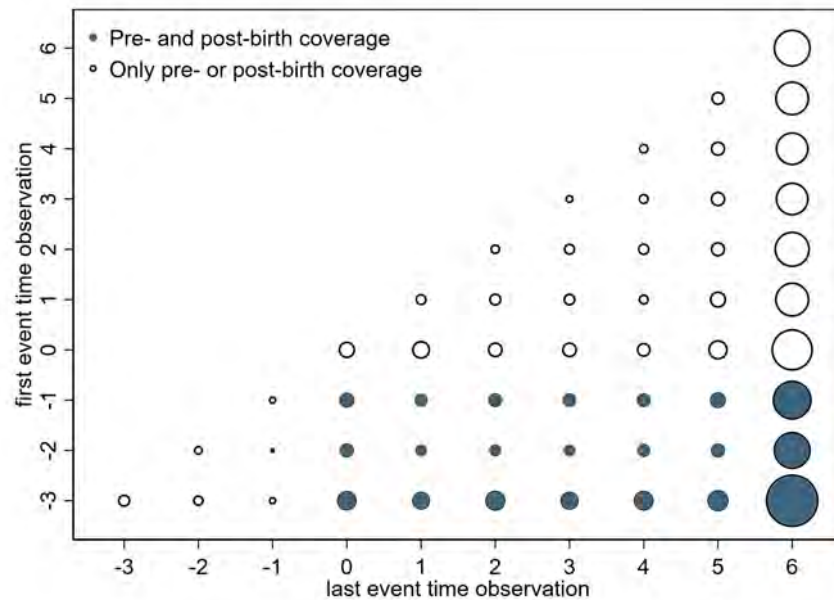
Deeply held gender norms may be difficult to influence in the short-run, but family policies such as expansions of day care or parental leave policies, may both facilitate maternal employment and have an impact on norms in the long-run³⁶ if trade-offs between family and career are reduced, thus providing a fruitful avenue to reduce child-related gender inequality.

³⁶E.g., Zoch and Schober (2018) find that day care expansion in West Germany is associated with less gender-traditional views. Dahl et al. (2014); Welteke and Wrohlich (2019) show spill-over effects of parental leaves decisions after policy reforms are implemented. Recognition of same-sex relationships in Europe has been accompanied by more positive attitudes towards sexual minorities (Aksoy et al., 2020).

7.A Appendix

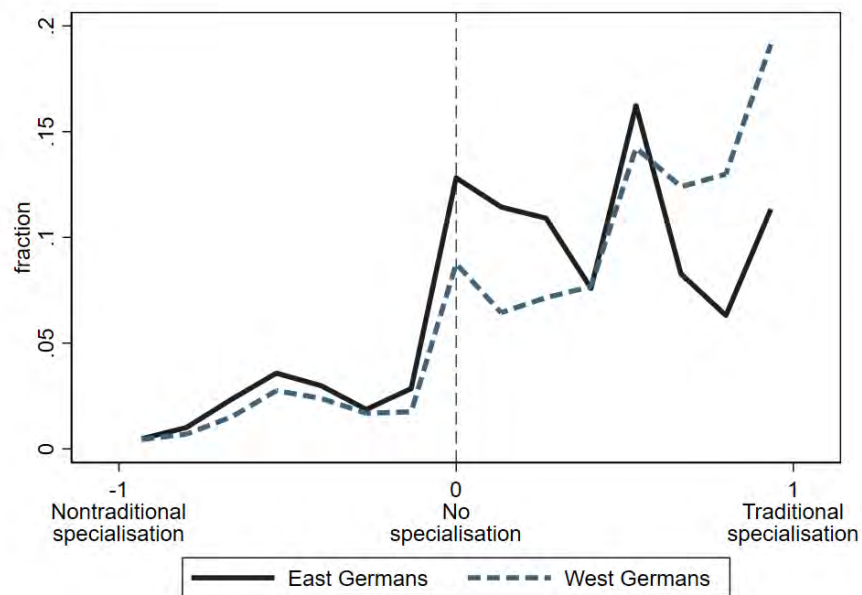
7.A.1 Figures

Figure 7.A.1: Observations by event time



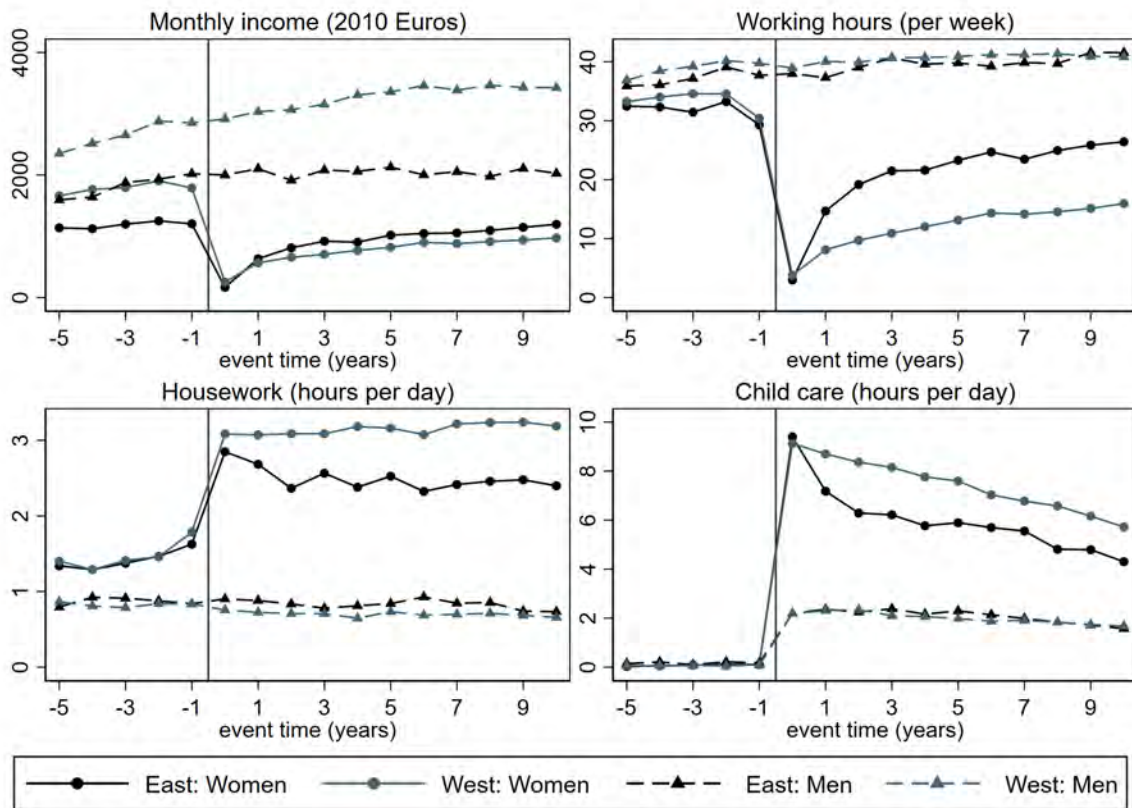
Note: Figure shows how observations in the event study estimates are distributed by their first and last event time (first birth) observation. Circle size is weighted by the number of observations in each combination. Blue coloured circles are observations that include both pre- and post-birth event time. Appendix Table 7.A.1 shows estimates for the full sample and for the sample with pre- and post-birth information. Source: SOEP v36

Figure 7.A.2: Specialisation index



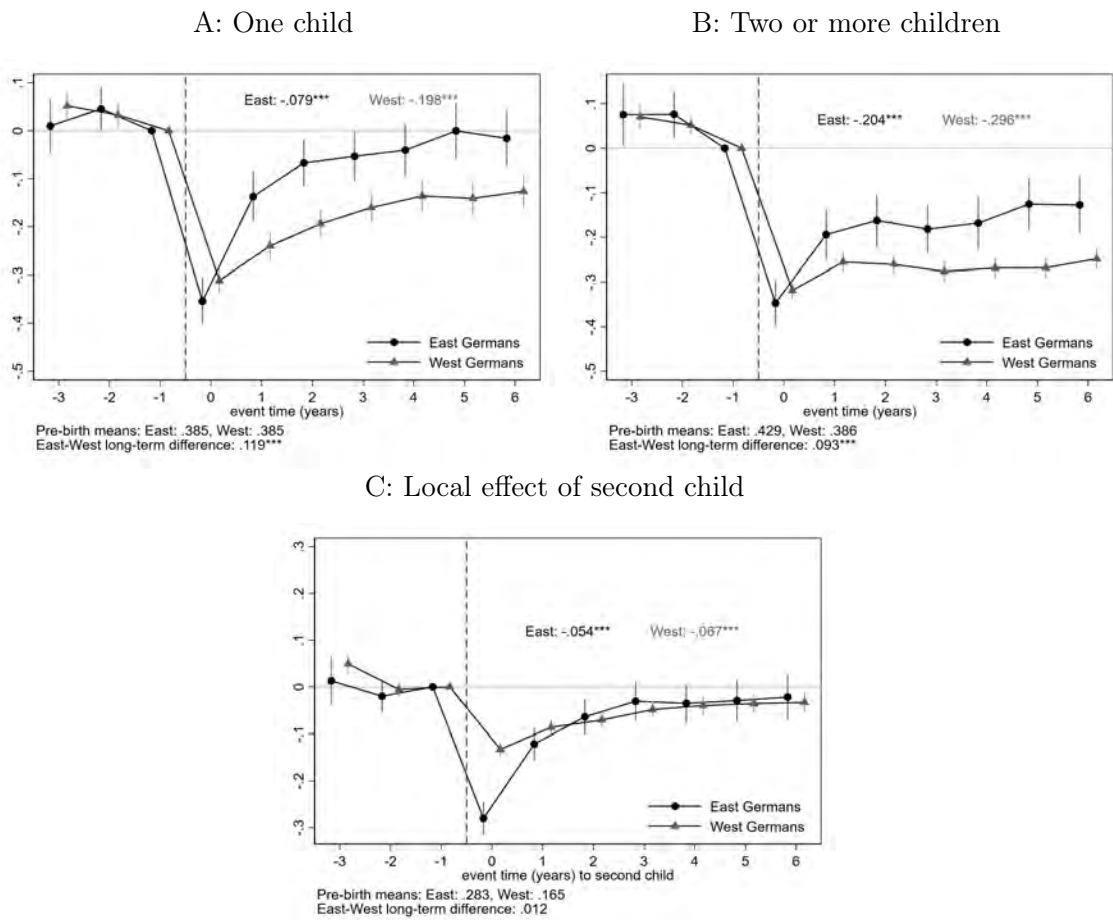
Note: Figure shows the distribution of a specialisation index proposed by Siminski and Yetsenga (2020). Traditional specialisation implies that the female partner is solely responsible for domestic work and the male partner for market work (vice versa for nontraditional specialisation). Distribution calculated in 15 bins of equal width. Source: SOEP v36

Figure 7.A.3: Average values by event time



Note: Figure shows average values of the respective variables by event time relative to the birth of first child. Sample covers 1990-2019. Source: SOEP v36

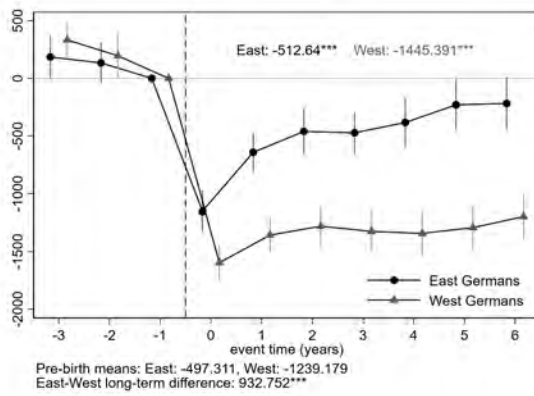
Figure 7.A.4: Impact of children on female income share - by number of children



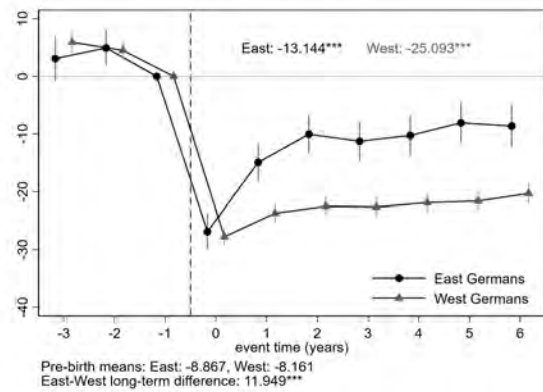
Notes: Panels A and B show event study estimates for the female income share differentiated by number of children. Panel C shows the local effect of having a second child. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.5: Impact of children on gender inequality in the labour market - gaps and discrete outcomes

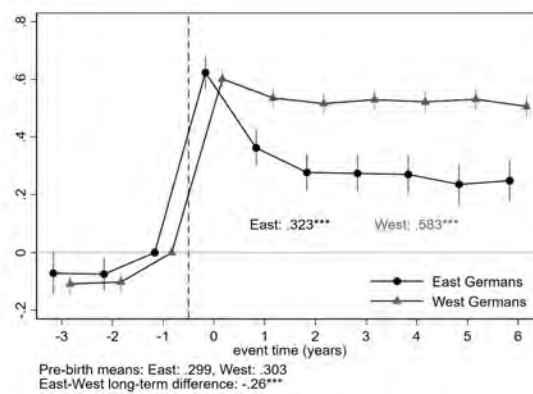
A: Gap in monthly labour income (2010 Euros)



B: Gap in weekly working hours

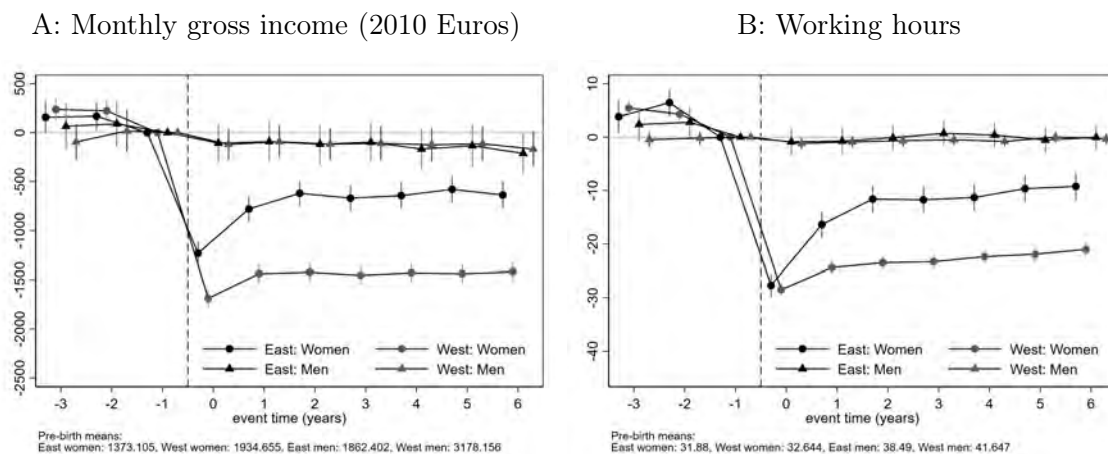


C: Main male breadwinner (0/1)



Notes: Panels A and B show gaps corresponding to shares shown in Figure 7.3. Main male breadwinner households are defined as such if the female income share is below $\frac{1}{3}$. See Figure 7.3 for other notes. Source: SOEP v36

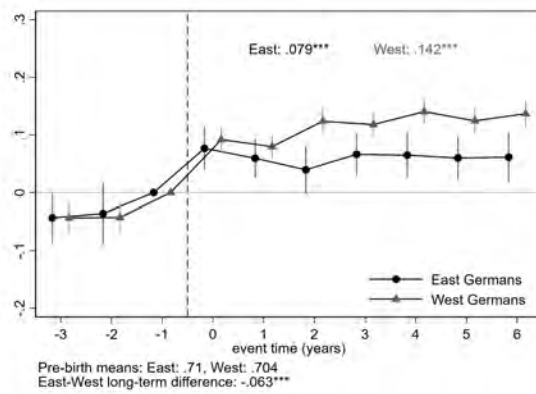
Figure 7.A.6: Impact of children on gender inequality in the labour market - individual level



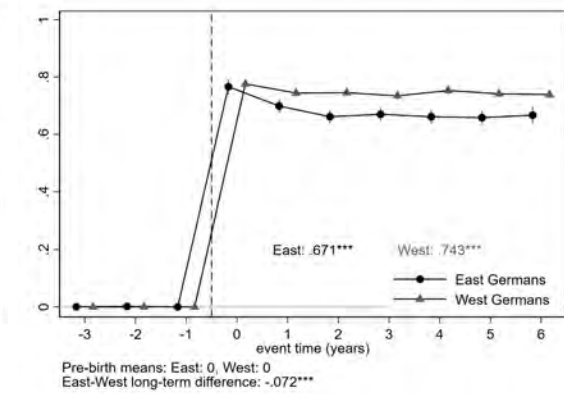
Notes: Panels A and B shows individual-level contributions to the shares depicted in Figure 7.3. Panel C shows the share that indicate currently being in parental leave. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.7: Impact of children on domestic gender inequality - weekly

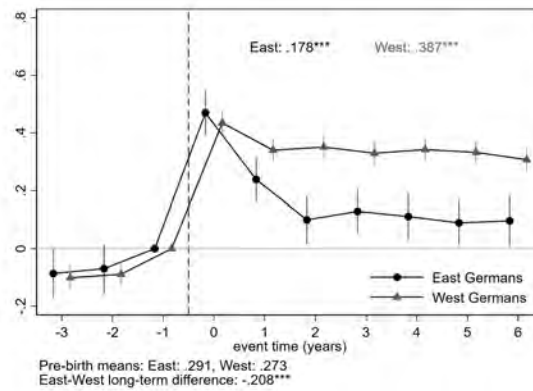
A: Housework share



B: Child care share

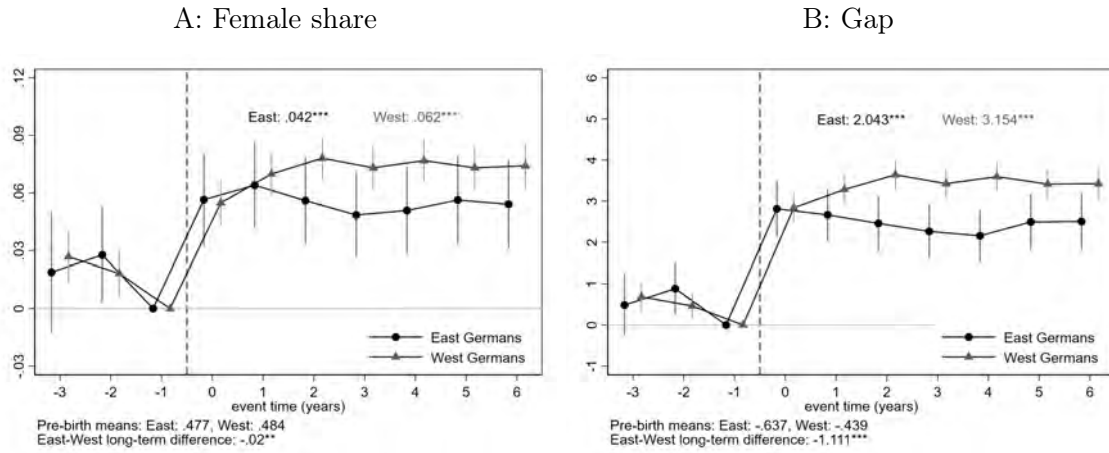


C: Specialisation index



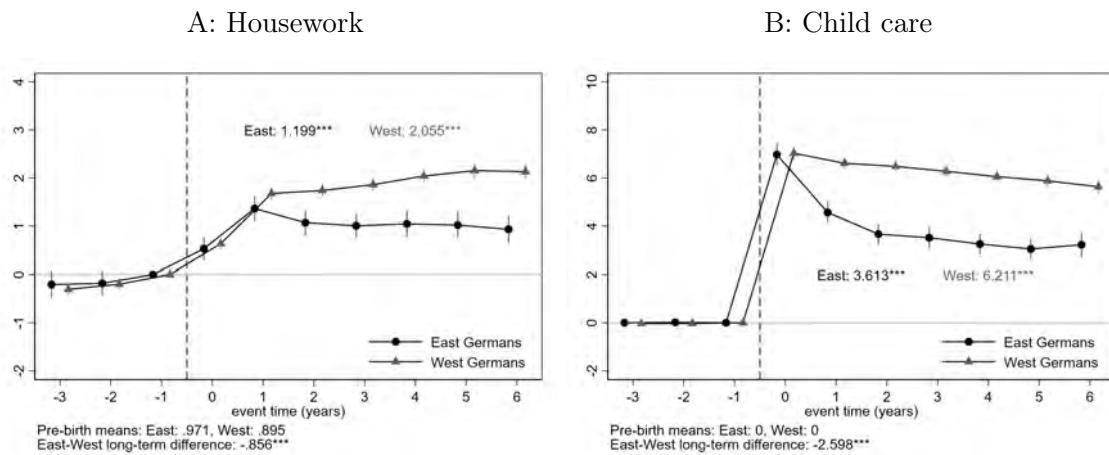
Notes: Figure shows event study estimates for the respective outcomes normalised to the pre-birth year ($t = -1$). In contrast to Figure 7.4, child care and housework information include weekend days, which reduces the sample size by about half. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.8: Impact of children on total hours of work



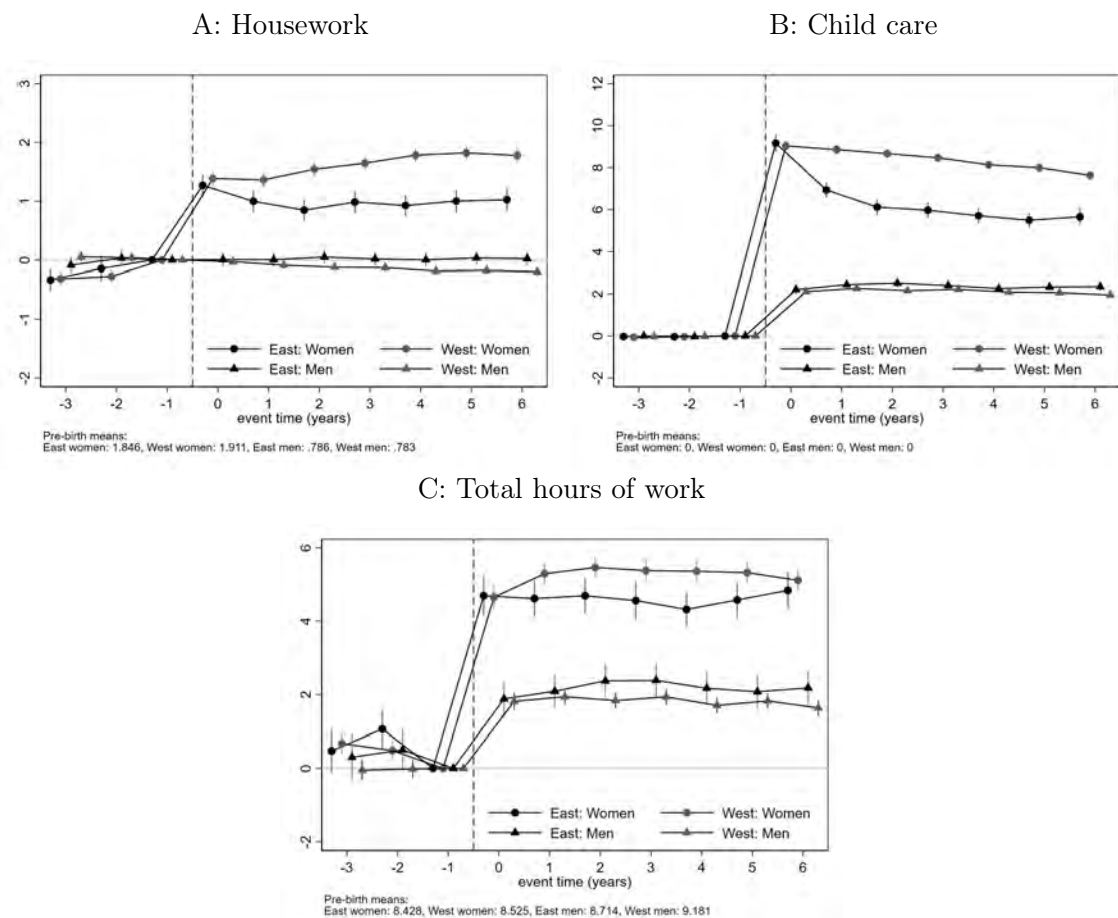
Notes: Figure shows event study estimates for total work, defined as paid work, housework and child care per weekday. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.9: Impact of children on domestic gender inequality - gaps



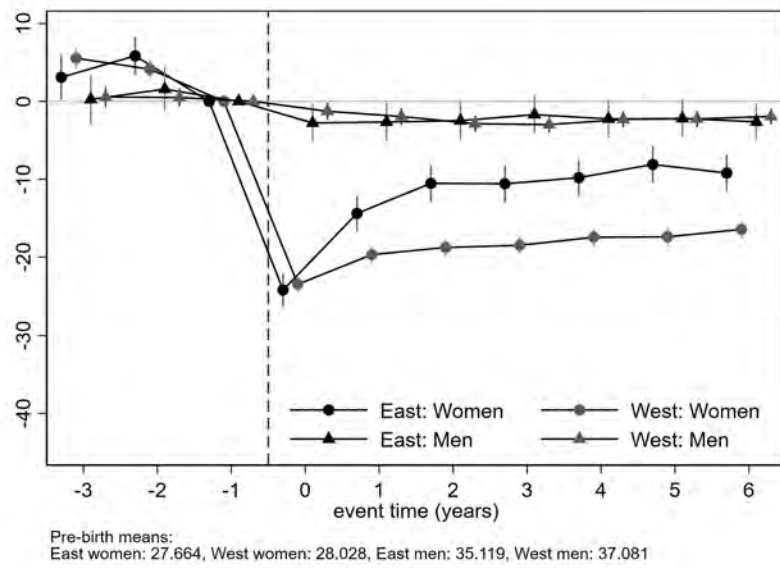
Notes: Figure shows gaps corresponding to shares presented in Figure 7.4. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.10: Impact of children on gender inequality in the labour market - individual level



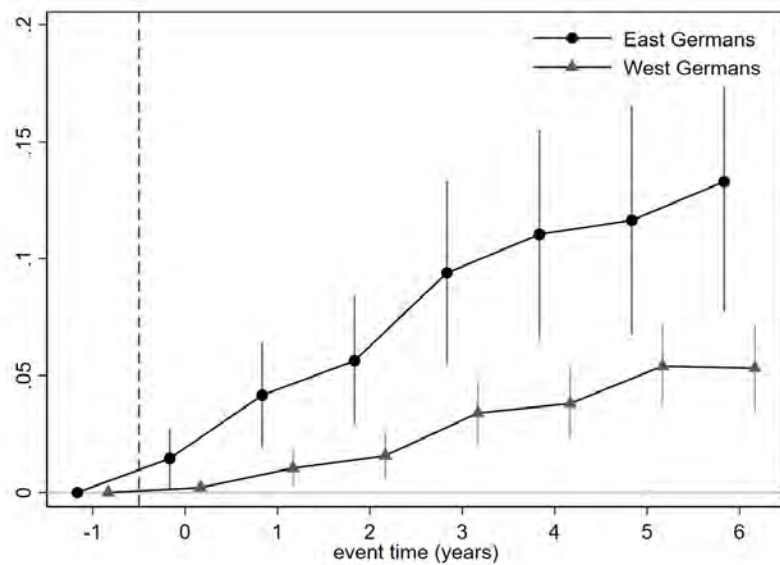
Notes: Panels show individual-level contributions to the shares depicted in Figure 7.4. See Figure 7.3 for other notes. Source: SOEP v36

Figure 7.A.11: Impact of children on desired working hours - individual level



Notes: See Figure 7.3 for notes. Source: SOEP v36

Figure 7.A.12: Couple dissolution



Notes: Figure shows event study estimates for couples to have dissolved. The sample is restricted to couples that existed in the year prior to child birth. Source: SOEP v36

7.A.2 Tables

Table 7.A.1: Long-run impacts of children - full sample and sample with pre- and post-birth observations

	East German couples		West German couples		East-West difference	
	(1)	(2)	(3)	(4)	(5)	(6)
<hr/>						
Female income share						
Long-term effect	-0.155*** (0.016)	-0.171*** (0.018)	-0.269*** (0.007)	-0.271*** (0.008)	0.114*** (0.017)	0.100*** (0.019)
<hr/>						
Female share of working hours						
Long-term effect	-0.158*** (0.015)	-0.177*** (0.017)	-0.292*** (0.007)	-0.296*** (0.008)	0.134*** (0.016)	0.119*** (0.018)
<hr/>						
Female housework share						
Long-term effect	0.077*** (0.013)	0.084*** (0.015)	0.158*** (0.007)	0.158*** (0.007)	-0.081*** (0.015)	-0.073*** (0.016)
<hr/>						
Female share of child care						
Long-term effect	0.709*** (0.008)	0.714*** (0.011)	0.789*** (0.003)	0.789*** (0.004)	-0.079*** (0.009)	-0.075*** (0.012)
<hr/>						
Specialisation index						
Long-term effect	0.189*** (0.024)	0.211*** (0.028)	0.411*** (0.012)	0.411*** (0.012)	-0.222*** (0.026)	-0.200*** (0.030)
<hr/>						
Age, survey year FEs	Y	Y	Y	Y	Y	Y
Pre- and post-birth observations		Y		Y		Y
Observations	4,088	2,513	12,552	8,898	16,640	11,411

Notes: Table shows long-run coefficients ($t = 1 - 6$) of the arrival of children on within-couple gender inequality. Columns (1), (3) and (5) are estimates shown in Figures 7.3 and 7.4. The odd-numbered columns restrict the sample to households with pre- and post-birth observations. Standard errors clustered at the couple-level in parentheses. Significance levels: * < 0.1 ** < 0.05 *** < 0.01. Source: SOEP v36

Table 7.A.2: Test of approximated control group

	Female share of							
	Monthly income		Working hours		Housework		Child care	
	True (1)	Appr. (2)	True (3)	Appr. (4)	True (5)	Appr. (6)	True (7)	Appr. (8)
East Germans:	-0.155*** (0.016)	-0.148*** (0.014)	-0.158*** (0.015)	-0.153*** (0.014)	0.077*** (0.013)	0.071*** (0.011)	0.711*** (0.007)	0.645*** (0.013)
Observations	3,890	4,240	3,724	4,050	4,088	4,426	4,618	4,156
West Germans:	-0.269*** (0.007)	-0.281*** (0.007)	-0.292*** (0.007)	-0.301*** (0.007)	0.158*** (0.007)	0.174*** (0.006)	0.789*** (0.003)	0.780*** (0.004)
Observations	12,332	13,779	11,792	13,165	12,552	13,989	13,947	12,380
East-West difference	0.114*** (0.017)	0.133*** (0.016)	0.134*** (0.016)	0.148*** (0.015)	-0.081*** (0.015)	-0.104*** (0.013)	-0.079*** (0.009)	-0.135*** (0.013)
Observations	16,222	18,019	15,516	17,215	16,640	18,415	15,592	16,536

Note: Table displays coefficients of (true) effects of having children based on equation (7.3) which uses pre- and post-birth information. Approximated coefficients are based on regressions of couples having children aged one to six with childless couples of a similar age range (5th to 95th percentile of age distribution of couples with children). Source: SOEP v36

Table 7.A.3: Couple-level sums for main outcomes

	Monthly gross income (2010 Euros) (1)	Working hours per week (2)	Housework on weekday (3)	Child care on weekday (4)
<i>Event time</i>				
-3	4,672.08 (2260.68)	76.18 (22.70)	2.41 (1.38)	0.00 (0.00)
-2	4,953.02 (2720.35)	77.55 (19.79)	2.45 (1.51)	0.00 (0.00)
-1	4,755.69 (3201.46)	73.06 (23.39)	2.62 (1.48)	0.00 (0.00)
0	3,141.66 (2002.70)	44.40 (18.03)	3.85 (2.13)	10.80 (5.31)
1	3,535.14 (2294.72)	50.10 (19.31)	3.82 (2.03)	10.64 (4.57)
2	3,598.40 (2037.78)	51.51 (19.99)	3.87 (2.01)	10.34 (4.48)
3	3,811.92 (2465.03)	53.70 (19.05)	3.95 (1.99)	9.99 (4.83)
4	4,030.49 (2864.86)	54.86 (19.68)	3.88 (2.04)	9.56 (4.73)
5	4,166.14 (2786.62)	56.17 (19.93)	3.92 (2.13)	9.34 (4.64)
6	4,260.06 (3219.49)	57.22 (20.66)	3.92 (2.05)	8.78 (4.64)
Observations	34,147	32,812	34,007	35,243

Note: Table shows the sum of the main outcomes for couples from which shares are calculated as dependent variables. Standard deviations in parentheses. Source: SOEP v36

Table 7.A.4: Impact of children - time-use data

	Female share of				
	Market work (1)	Housework (all) (2)	Housework (rout.) (3)	Child care (4)	Specialisation (5)
GDR	-0.057* (0.030)	-0.005 (0.036)	-0.010 (0.034)	0.875*** (0.013)	0.100* (0.055)
Observations	661	678	678	649	652
East Germany	-0.143*** (0.036)	0.033 (0.028)	0.047 (0.029)	0.747*** (0.018)	0.255*** (0.053)
Observations	750	875	874	864	755
West Germany	-0.207*** (0.017)	0.104*** (0.014)	0.106*** (0.014)	0.748*** (0.009)	0.329*** (0.027)
Observations	2,227	2,640	2,637	2,627	2,207

Note: Table displays conditional differences in time use of households with and without children. The sample of households with children have at least one child below the age of 6, and the sample of households without children are set to be in the 5th to 95th female age percentile of those with children. All estimates include survey wave and age FEs and additionally control for education and marital status. Sources: Time-Use survey of the GDR (1985, 1990) and German Time-Use Study (1991/92, 2001/02, 2012/13).

7.A.3 Robustness

Table 7.A.5: Estimates based on current location

	Female share of				
	Income (1)	Hours (2)	Housework (3)	Child care (4)	Specialisation (5)
East Germany	-0.147*** (0.015)	-0.155*** (0.015)	0.081*** (0.012)	0.708*** (0.007)	0.197*** (0.022)
Observations	4,087	3,907	4,332	4,838	3,682
West Germany	-0.263*** (0.006)	-0.284*** (0.006)	0.153*** (0.006)	0.778*** (0.003)	0.399*** (0.010)
Observations	17,173	16,518	17,634	18,502	15,063
East Germans in East Germany	-0.154*** (0.016)	-0.158*** (0.016)	0.072*** (0.014)	0.708*** (0.009)	0.185*** (0.025)
Observations	3,429	3,278	3,620	3,465	3,100
East Germans in West Germany	-0.162*** (0.053)	-0.157*** (0.047)	0.096** (0.041)	0.720*** (0.022)	0.185** (0.079)
Observations	461	446	468	437	408
West Germans in West Germany	-0.271*** (0.007)	-0.294*** (0.007)	0.157*** (0.007)	0.789*** (0.003)	0.411*** (0.012)
Observations	12,226	11,697	12,443	11,590	10,722

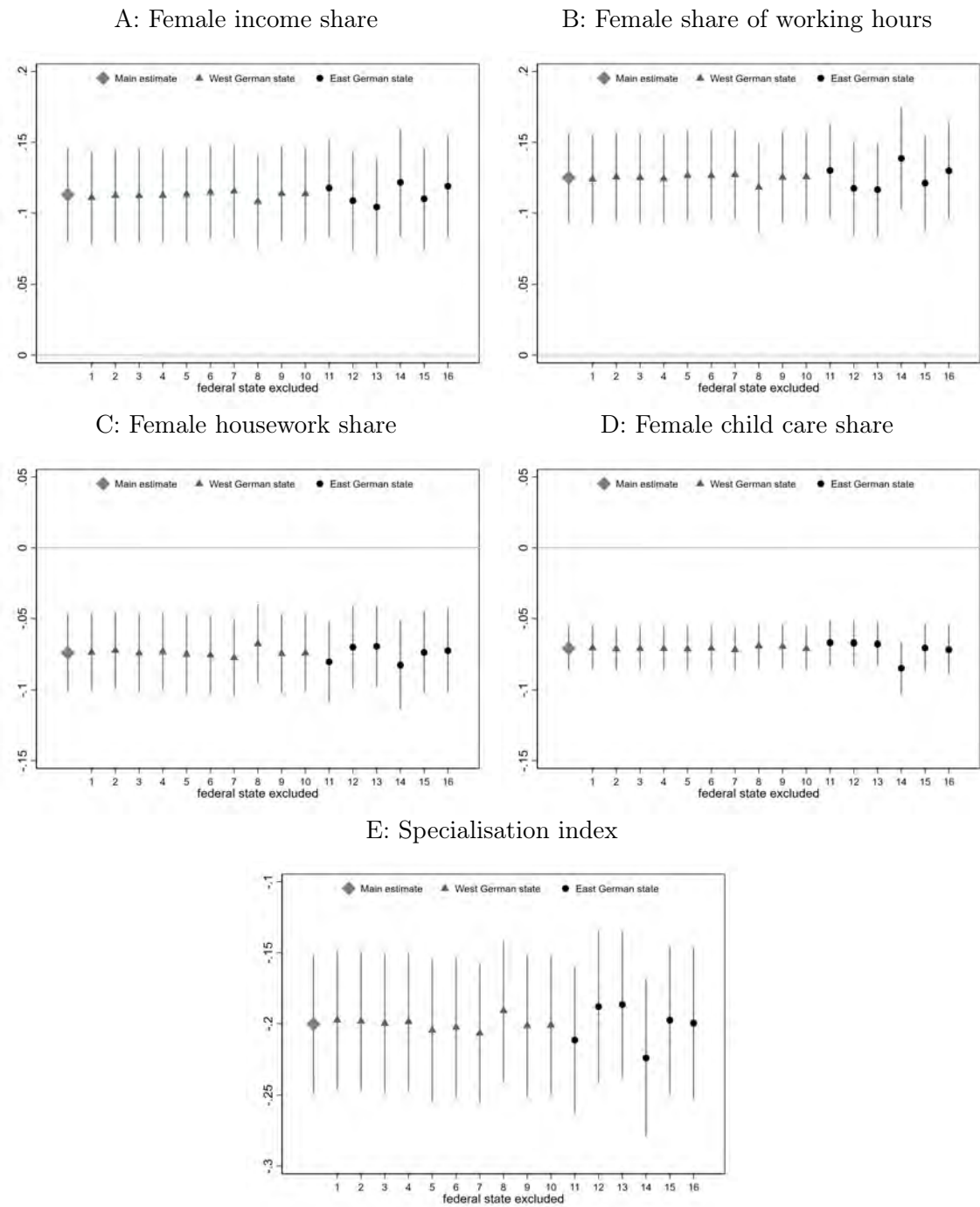
Note: Table shows coefficients for the long-term effect of children as in Table 7.3. East and West German couples are defined by their 1989 location. East and West Germany related to the current location with a distinction made between East and West Berlin. The two upper rows are only based on current location, the bottom three rows distinguish between location and origin of couples. Source: SOEP v36

Table 7.A.6: Both partners born in Germany

	Female share of				
	Income (1)	Hours (2)	Housework (3)	Child care (4)	Specialisation
East Germans	-0.156*** (0.016)	-0.160*** (0.015)	0.078*** (0.013)	0.710*** (0.008)	0.192*** (0.024)
Observations	3,814	3,655	4,009	3,827	3,442
West Germans	-0.278*** (0.008)	-0.301*** (0.007)	0.160*** (0.008)	0.791*** (0.004)	0.427*** (0.012)
Observations	10,460	10,050	10,607	9,851	9,197

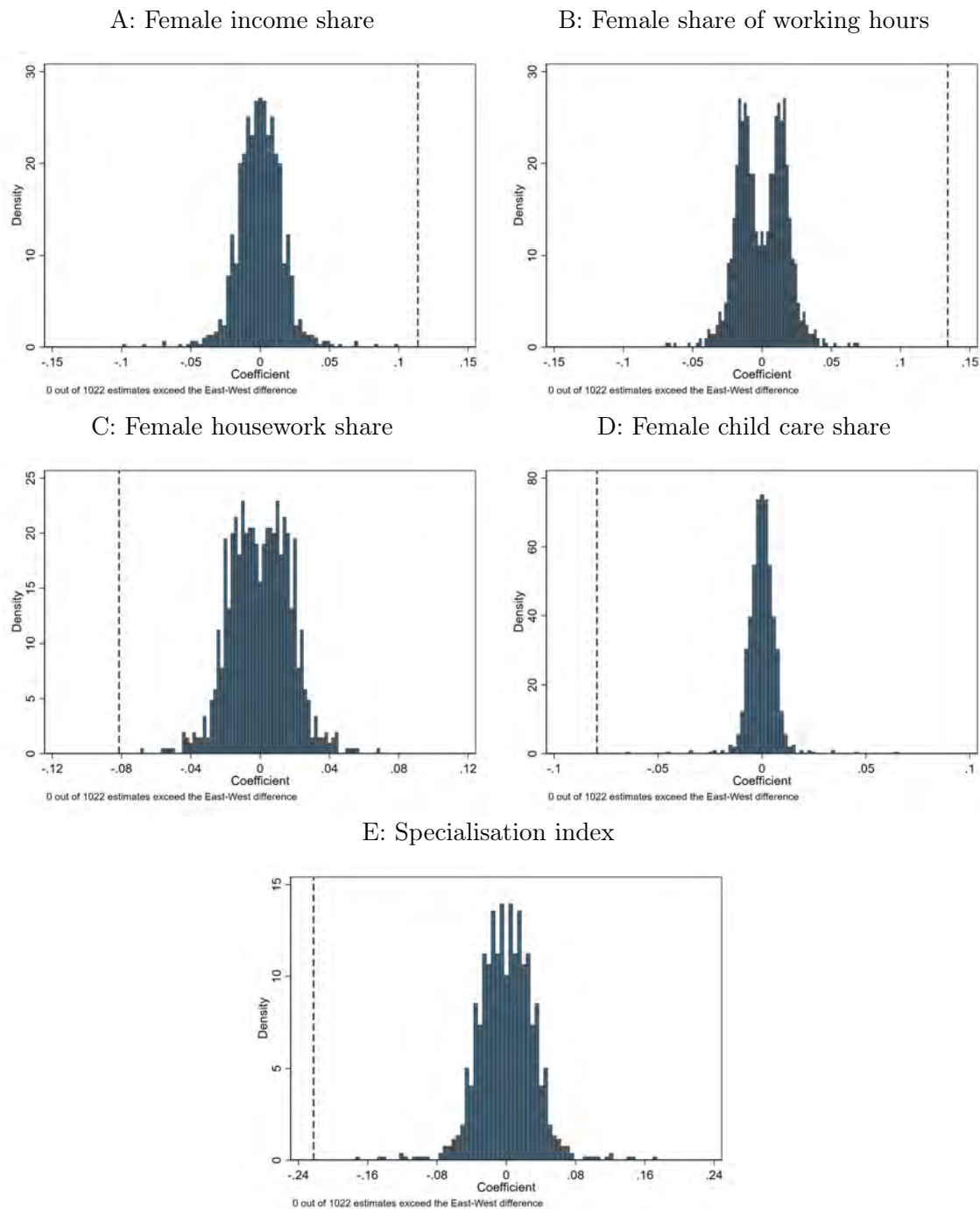
Note: Table shows coefficients for the long-term effect of children as in Table 7.3 with the estimation sample restricted to couples where both partners were born in Germany. Main estimates in contrast are based on the 1989 location with no further restriction on birth place. Source: SOEP v36

Figure 7.A.13: East-West long-run difference with states excluded



Notes: Figure shows East-West long-run differences of the arrival of children on within-couple gender inequality with states subsequently dropped in the estimation. The grey main estimates refer to estimates based on the current location of couples. List of states: 1 Schleswig-Holstein 2 Hamburg 3 Lower Saxony 4 Bremen 5 North Rhine-Westphalia 6 Hesse 7 Rhineland-Palatinate 8 Baden-Württemberg 9 Bavaria 10 Saarland 11 Berlin 12 Brandenburg 13 Mecklenburg-Vorpommern 14 Saxony 15 Saxony-Anhalt 16 Thuringia. Source: SOEP v36

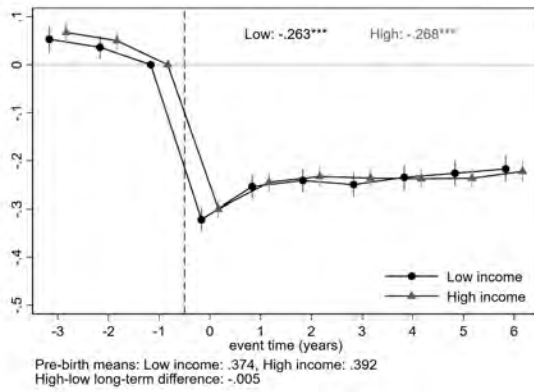
Figure 7.A.14: Placebo treatment assignment



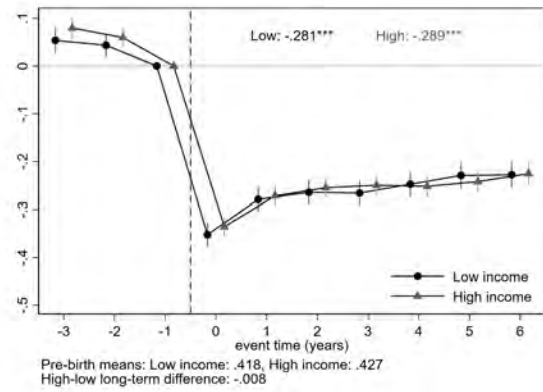
Note: Figure shows the distribution of estimates for long-run differences between two groups of federal states. The histogram displays estimates based on all 1022 possible combinations of West German states into two groups. The dashed lines denote the East-West difference based on the federal state of residence. Figure note indicates how many of the placebo estimates exceeded the coefficient for the East-West difference. Source: SOEP v36

Figure 7.A.15: West German counties split by income per capita

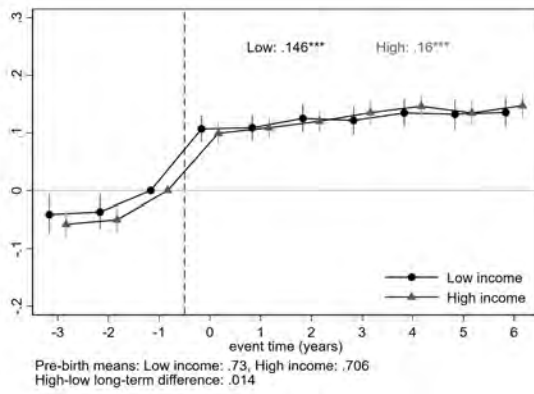
A: Female income share



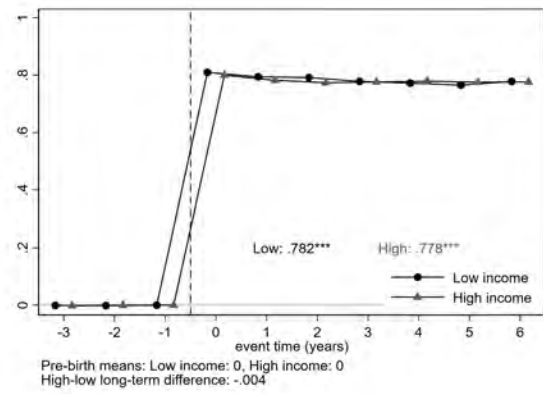
B: Female share of working hours



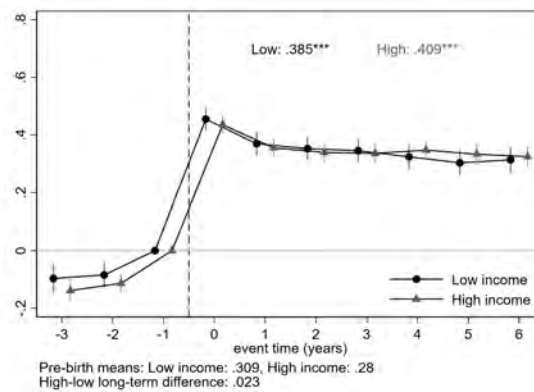
C: Female housework share



D: Female child care share



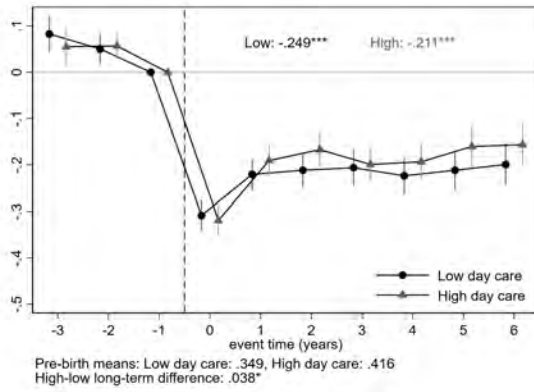
E: Specialisation index



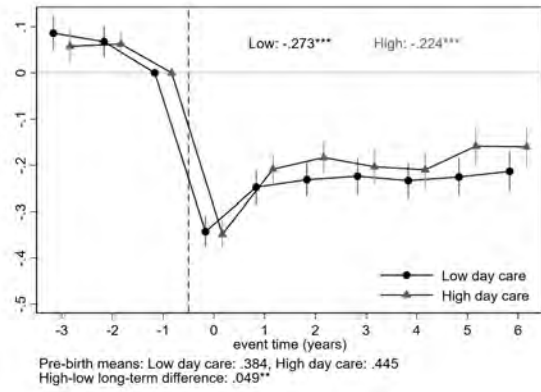
Note: Figure shows event study estimates corresponding to Figures 7.3 and 7.4. The sample is restricted to couples living in West Germany and the sample is split by GDP per capita as of 2008. See Figure 7.3 for other notes. Source: SOEP v36 and Destatis

Figure 7.A.16: West German counties split by day care availability

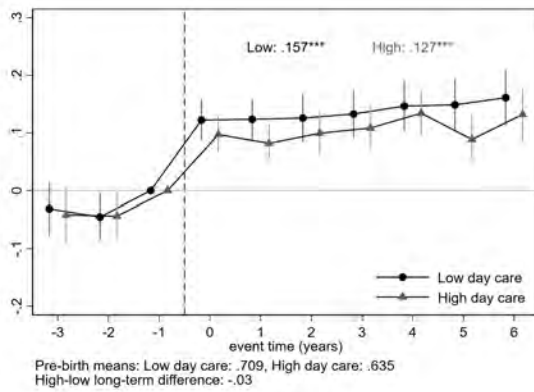
A: Female income share



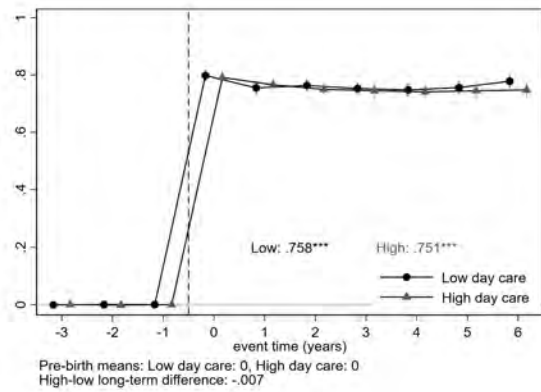
B: Female share of working hours



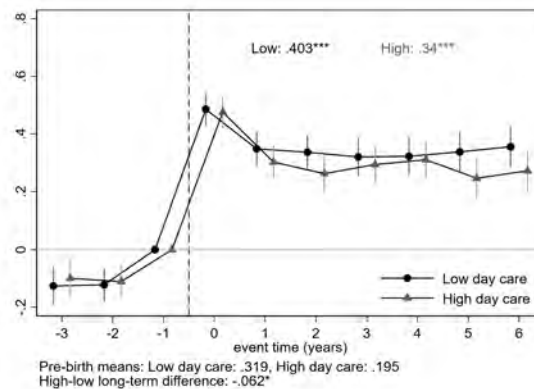
C: Female housework share



D: Female child care share

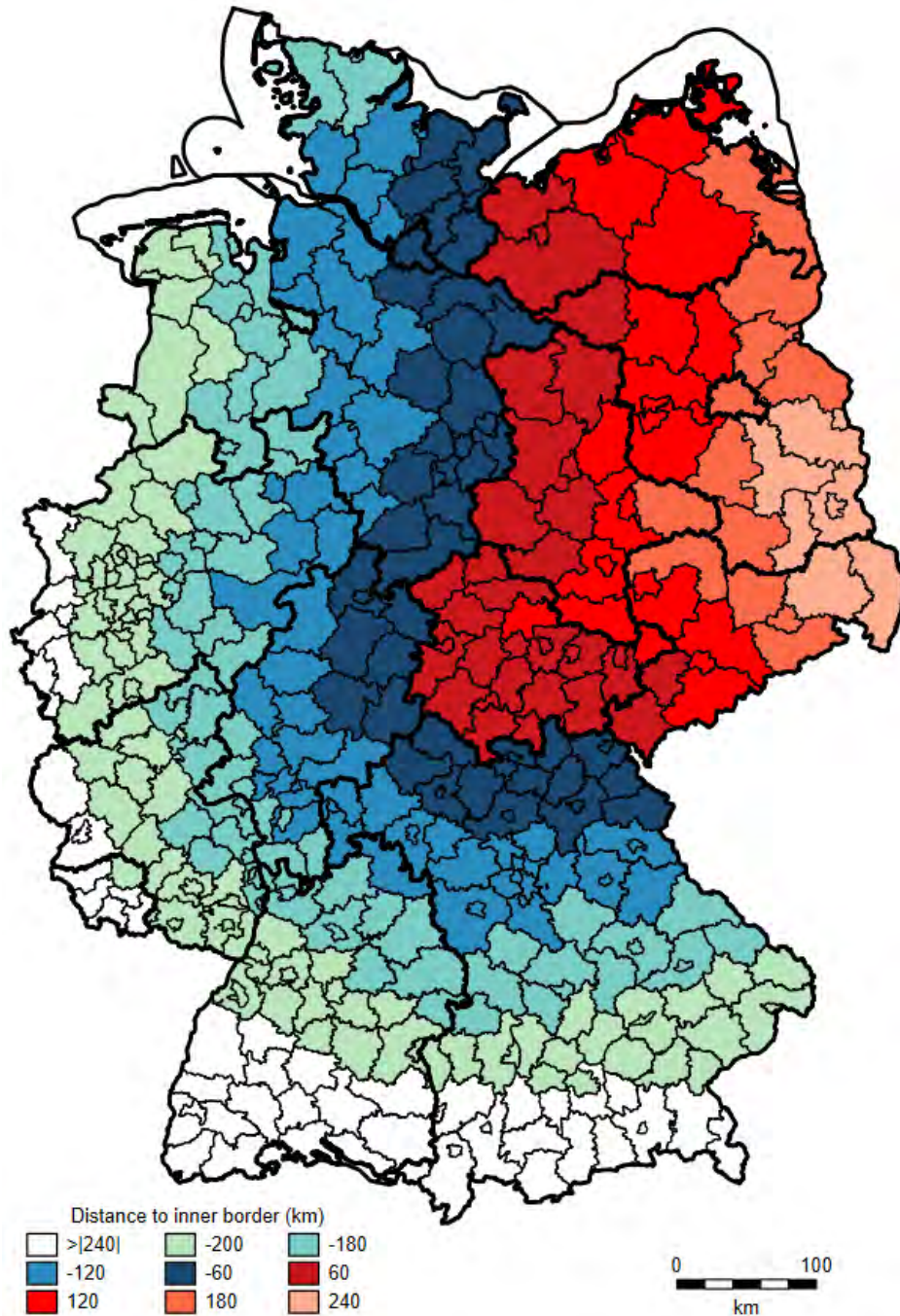


E: Specialisation index



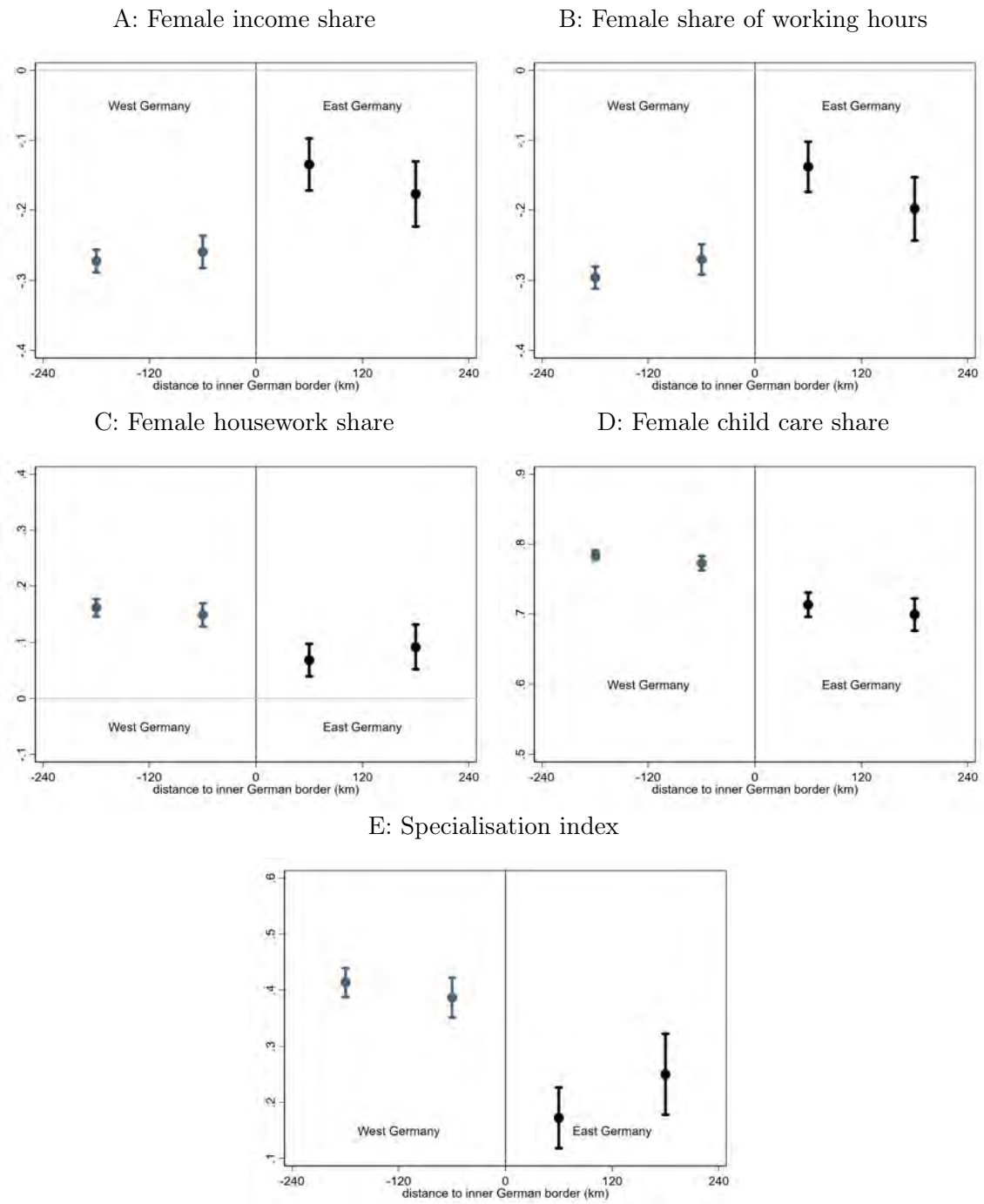
Note: Figure shows event study estimates corresponding to Figures 7.3 and 7.4. The sample is restricted to couples living in West Germany and the sample is split at the county-level the share of children under the age of three enrolled in day care. Estimates based on births occurring between 2004 and 2019. See Figure 7.3 for other notes. Source: SOEP v36 and Destatis

Figure 7.A.17: German counties and distance to inner border



Note: Map shows contemporaneous German counties and their geodesic closest distance to the former inner German border. Distances are calculated from counties' centroids. Red shades indicate East German states, blue shades West German states. Shapefiles from the Federal Agency for Cartography and Geodesy

Figure 7.A.18: Spatial RD



Note: Figure shows long-run estimates of the arrival of children on within-couple gender inequality by distance to the inner German border. Distances are based on the current county of residence and are calculated from counties' centroids to the closest border point. Estimates are calculated in 120km bins to the border. Source: SOEP v36

7.A.4 Comparison of time-use data from diary data and SOEP

Time-use diary data is generally considered to contain less measurement error than survey data based on retrospective questions when it comes to accurately depicting individuals' time spent on various activities (Marini and Shelton, 1993). Diary data is commonly recorded throughout the day (or after a day) in small time slots. The German Time-Use Survey asks participants to record their days in five- to ten-minute slots. In the SOEP questions are asked for a 'typical' weekday or weekend day and respondents may only indicate full hours, automatically leading to some activities being under- or overreported. The literature has found that especially unpaid work, which is usually carried out in irregular intervals, is difficult for respondents to accurately estimate (Kitterød et al., 2005). The precision of time-diary data comes at the expense of less background information available in such data on individuals, smaller sample sizes and the lack of a panel structure (in Germany, as in most other countries). Longitudinal data is essential for most analyses in this paper, making it desirable to be able to use information from the SOEP for some analyses.

In this Appendix section I compare averages obtained from the SOEP and from the German Time-Use Survey, to verify the usage of survey data. In a first step some restrictions have to be imposed to make the samples more comparable. SOEP data is restricted to the same years as the three waves of the time-use survey (1991/92, 2001/02, 2012/13). In both data sets, only information from weekdays is used. Additionally I focus on the main group of interest; couples with exactly one child below the age of six.³⁷

The time-use survey contains detailed 3-digit activities, e.g. the 3-digit category 'baking' belongs to the 2-digit category 'preparation of meals' of the 1-digit category 'housework'. In comparison the SOEP questionnaires ask for the time spent on housework, but specifying that it refers to washing, cooking and cleaning. Simply using the 1-digit category housework from the time-use data leads to large differences between the data sets with on average 50% more time spent on housework in the time-use survey. To ensure better overlap between the housework information, a narrower definition of housework consisting of the 2-digit categories 'preparation of meals', 'maintenance and cleaning of the house or flat', and 'fixing textiles' is defined from the time-use survey, which mirrors the definition in the SOEP.

Similarly, for child care, using the 1-digit category of the time-use survey initially leads to large differences with almost 200% more time spent on this in the SOEP.

³⁷The reason for this restriction is that the time-use data only contains information on the age of the youngest child in the household. In the event study estimates in this paper, in contrast, the event time relates to the birth of the first child (i.e. the oldest).

The retrospective questions in the SOEP generally allow for parallel activities and combining all the different activities elicited often adds up to more than 24 hours per day, whereas primary activities in the time-use survey are by definition mutually exclusive. Child care in the time-use survey consists of specific activities with the child, e.g. playing with or reading to the child. Besides the activities, the time-use survey also contains indicators on whether the child was present at any time (Jessen et al., 2020b, use the terms ‘parenting activities’ and ‘time with child’ to differentiate between those). As parents of young children will still be interacting with the children and be somewhat constrained by their presence, it is not an unreasonable to assume that a general question for ‘time spent on child care’ will be interpreted this way.

Appendix Table 7.A.7 shows a comparison of time spent on housework and child care using the definitions described above. Panel A shows averages from the SOEP, and Panel B from the time-use survey. Both panels differentiate by survey wave, and displays results separately for women, men, the female share, and by location in East and West Germany. While the values are not perfectly in line, perhaps not too surprising given different sampling, and retrospective questions vs. time diary, it is still apparent that results from the two data sets are generally comparable and differences between different groups (by region or survey year) also tend to point in the same direction. This reassures that time-use information from the SOEP can be used reliably in the analysis.

Table 7.A.7: Time use comparison

	Housework			Child care		
	1991/1992 (1)	2001/2002 (2)	2012/2013 (3)	1991/1992 (4)	2001/2002 (5)	2012/2013 (6)
Panel A: SOEP						
East						
Women	2.60 (1.43)	2.56 (1.50)	1.68 (1.00)	4.91 (3.64)	6.16 (4.03)	5.08 (3.27)
Men	0.53 (0.64)	0.78 (0.71)	0.99 (0.86)	1.98 (1.96)	2.09 (1.30)	2.36 (1.85)
Female share	0.81 (0.20)	0.75 (0.21)	0.71 (0.24)	0.72 (0.23)	0.71 (0.17)	0.71 (0.20)
Observations	236	261	253	234	259	250
West						
Women	3.42 (1.89)	3.00 (1.98)	1.93 (1.11)	7.24 (3.42)	7.90 (4.05)	7.40 (4.65)
Men	0.49 (0.66)	0.69 (1.03)	0.78 (0.75)	1.80 (1.61)	2.16 (1.84)	2.60 (2.65)
Female share	0.86 (0.20)	0.82 (0.22)	0.73 (0.26)	0.78 (0.18)	0.77 (0.17)	0.71 (0.23)
Observations	737	957	693	725	932	682
Panel B: Time-Use Survey						
East						
Women	2.58 (1.70)	1.71 (1.55)	1.61 (1.29)	5.93 (3.60)	6.77 (4.05)	6.02 (3.28)
Men	0.83 (1.02)	0.57 (0.56)	0.49 (0.61)	2.95 (3.05)	3.41 (2.92)	3.06 (2.92)
Female share	0.62 (0.26)	0.53 (0.27)	0.61 (0.29)	0.75 (0.26)	0.75 (0.29)	0.75 (0.27)
Observations	872	100	212	872	100	212
West						
Women	3.73 (1.80)	2.47 (1.67)	1.99 (1.52)	8.88 (3.64)	7.13 (3.05)	6.74 (3.41)
Men	0.65 (0.93)	0.69 (1.14)	0.57 (0.83)	3.04 (2.59)	3.25 (2.92)	2.91 (2.47)
Female share	0.70 (0.22)	0.67 (0.26)	0.65 (0.27)	0.77 (0.24)	0.76 (0.25)	0.73 (0.27)
Observations	2,362	408	630	2,362	408	630

Note: Table compares average values for time spent on housework and on child care using SOEP survey data and the German Time-Use Survey. The SOEP comparison sample is restricted to the same years as the time-use data. Averages refer to weekdays and observations are restricted to couples with a child between the age of one and six years. Source: SOEP v36 and German Time-Use Survey

CHAPTER 8

Conclusion

This dissertation comprises of six chapters examining how gender and socio-economic inequality is impacted by family policies in the German context. Two focus points are what unintended effects these family policies may have and what the legacy of differential family policies in the GDR and the FRG have on gender inequality in reunified Germany. This final chapter draws a conclusion by briefly discussing policy implications, limitations of the studies¹ and fruitful avenues for future research.

Chapters 2 and 3 demonstrate that when designing parental leave policies the perspective of firms needs to be considered. If firms have to cover (part of) wage replacements directly, this disincentives them to hire potential mothers (**Chapter 2**). But even if their costs of motherhood (or parenthood) only come through the channel of having to replace a mother (or father) going on leave, these costs can be large with potentially detrimental effects for firms and potential mothers (**Chapter 3**). Firms do not replace all mothers going on leave, which then leads to less value added in those firms. Parental leave policies have to strike a balance between helping to reconcile work and family life both in the short- and long-run. Generous parental leave policies help families in the short-run but can have negative effects in the long-run.

A first policy implication is that it is crucial to identify factors that asymmetrically make young women more costly from a firm-perspective as this would have negative effects on women's labour market trajectories compared to men's. Currently, a large share of paid parental leave in Germany is taken by mothers despite the 2007 paid parental leave reform that incentivised more leave taking by fathers.² If more gender-

¹A more thorough discussion of this is found in the chapters as this overarching conclusion cannot fully cover this. A natural limit for all empirical work is that results are to some degree context-specific, and that the findings rely on the validity of the identifying assumptions underlying the empirical approaches.

²The reform was partly successful in the regard, as the share of fathers taking leave has strongly increased. But on average fathers still take much less leave than mothers; in 2020, fathers took 24.8% of all paid parental leave months (Destatis, 2021).

equal leave taking was encouraged, this would reduce negative effects on young women compared to young men. However, this would not eliminate statistical discrimination against potential mothers, but only reduce it and lead to *potential fathers* being affected as well. Alternatively, policy makers could reduce statistical discrimination against *potential parents*, for example, by compensating firms for birth events and related worker-absences.

A limitation of the chapters is that **Chapters 2 and 3** focus on firms of specific sizes (firms with more than 30 and less than 50 employees, respectively) for the sake of internal validity. Finding convincing research designs to study the effects on firms of other sizes³ and identifying effects of several birth-related absences within one firm are likely to be insightful fields for future work. Another important aspect these chapters could not examine is the effects of increasing utilisation of parental leave taking of fathers and compensatory labour supply adjustments within couples, as households and fathers cannot be identified in the German data. Similarly, we lack information on hourly wages which would also be revealing to consider as an adjustment mechanism by firms.

Many findings of these chapters will be applicable to other contexts as well as reactions from firms to parental-leave related incentives could be comparable, still our analyses come with the common limits to generalisability for other settings such as labour market institutions or gender norms. Studying the effects of birth-related absences in other institutional contexts seems to be a promising research field, as besides recent studies from Scandinavia (Gallen, 2019; Ginja et al., 2020b) evidence on firms and parental leave is lacking. This could hopefully cover settings with other lengths or generosity of paid parental leave and parental leave reforms at other margins where results are expected to differ.

Chapter 4 analyses enrolment gaps by parental education and migrant background in the German day care system. We find that differences in demand only play a small role, and reducing shortages or parental fees only closes the gap by parental education but not by migrant background. Policy recommendations to close the gap by parental education are relatively straightforward; reducing local shortages by increasing supply and abolishing fees (or increasing the progressiveness) would be effective tools. As small gaps in demand also exist, campaigns to boost demand could further reduce gaps.

For the gap by migrant background however, these policies would be less effective. Quality concerns appear to play a role, ensuring that preferences of parents from cul-

³We note that findings of **Chapter 3** are not driven by the size of firms, but by the availability of internal substitutes (workers in the same occupation). The effects may accordingly hold for larger firms with smaller workgroups as well.

turally diverse backgrounds are met would thus be important. Additionally, as a large share of the gap by migrant background is unexplained, discrimination could play a role. Training staff to be more sensitive to explicit and implicit biases and an impartial application system with transparent selection criteria are tools to reduce potential impacts of discrimination. It is also possible that the migrant gap would be reduced if shortages are reduced below levels that we observe in our data.

The understanding of discrimination in the context of day care, as well as in other non-market contexts, is still limited. We cannot rule this out as a residual explanation, but we are unable to explicitly test for this. Field experiments in form of audit studies could help fill this gap. Another important aspect we are unable to assess is differences in search effectiveness which is likely to be closely associated with social capital. I.e., some of the observed enrolment gaps by migrant background—where demand is similar to that of natives—could be due to German-born parents having an edge in the search and application process. This is arguably only relevant in the context of higher shortages when parents are competing for limited spots. Finally, providing parents with information and assisting them in the application process in a randomised setting could yield insights into the importance of those barriers (see Hermes et al., 2021).

The findings of the chapter are highly relevant for the German context, but day care systems vary strongly internationally in many regards, e.g., in terms of the application system, selection criteria, parental fees or quality standards. The degree of institutional similarity and of the population composition will determine to what extent these findings can be applied to other contexts, which must be assessed on a case-by-case basis. Naturally the results will be more transferable to universal, highly subsidised systems with insufficient provision. Future work could aim at better understanding the gaps in other countries and propose solutions to close them to ensure that children regardless of their background have the same chance to attend day care.

Chapter 5 estimates how parenting activities differ when their children are enrolled in day care. The findings highlight the importance of the home environment and, besides focusing on improving access to and quality of day care, policy makers could consider measures to further strengthen the home environment in order to improve development effects of day care on children. This could be done by, e.g., facilitating better interaction between day care staff to better coordinate complementarities between parental care and day care, or through home-visiting programmes aimed at improving the home environment (e.g., Doyle et al., 2017).

An obvious limitation of the chapter is that the analysis is correlational. As quasi-experimental approaches would be based on unreasonable assumptions, we instead include a rich set of covariates and present bounds for our estimates. We argue that

the estimates are reasonable proxies for causal effects, but of course a credible (quasi-) experimental approach would be an improvement. The data prevent us from doing this (the German Time-Use survey is only conducted once every decade⁴ and does not contain fine-grained regional information, making it ill-suited for a quasi-experimental approach), but in other settings researchers may find convincing research designs to estimate causal effects and building on our framework of direct and indirect effects. Availability of time-use diary data containing detailed activities by time of day and information on children's day care enrolment is a fundamental requirement. This chapter aims to understand the mechanisms behind the identified effects by looking at non-parenting activities and through this give a comprehensive picture of which activities parents are trading off. An advancement on this would be to develop a theoretical model to better understand parents' reactions and to run policy-simulations.

In **Chapter 6** we examine the relation between market and household work in the GDR and look at differences between East and West Germany. Women in East Germany contribute more time to market work and less to housework, and average within-couple housework gaps are much smaller in East Germany. Using a decomposition analysis, we find that housework gender gaps are almost identical once individual time constraints are accounted for. Gender inequality in housework was only reduced due to a mechanical effect, but conditional on own working time, East Germany is not found to be more gender equal regarding domestic work than West Germany.

Accordingly, policies targeting gender inequality in the labour market only impact inequality in non-paid work to a limited degree; in households with both partners working full-time, women are responsible for more than 60% of housework. If policy makers want to achieve more equal distributions of unpaid domestic work as well, other policies are required. E.g., more parental leave taking of fathers has been shown to lead to long-run effects in fathers' housework and child care contributions (Farré and González, 2019; Tamm, 2019). This chapter also has implications on models of household bargaining, which typically assume that individual housework contributions are (partly) determined by one's own and the partner's labour supply. We find that individual housework contributions are orthogonal to partner's labour supply (for both men and women).

As this chapter speaks to models of household bargaining, extending such models to explain our findings would be a fruitful addition to this work, which we aim to do in a follow-up project. A downside of our unique data is that it does not contain information on wages, an analysis with time-use data including wages would be intriguing as this

⁴The survey data we are using in this chapter are on an annual level, but the time-use information in this data is less detailed.

may be a more informative indicator of bargaining power. Finally, we note that more cross-country evidence on the reaction function between housework and market work would be interesting to assess the universality of our findings.

In the final **Chapter 7**, I examine how socialisation impacts within-couple gender inequality. Child penalties are found to be much smaller among East German couples (compared to West German ones), who were exposed to more gender egalitarian policies and norms in their childhood. As a large share of gender inequality is linked to children, it is important to understand determinants of child penalties. In a cross-country comparison, child penalties have been found to differ strongly (Kleven et al., 2019a). The setting of Germany's division and reunification allows to separate the effects of different culture or socialisation, and the institutional environment. This is comparable to the epidemiological approach where the children of different groups of immigrants living in the same country are studied (Fernández, 2011), but in the German context differently socialised individuals are arguably more similar in terms of social capital.

Social norms are often slow to change, but this work shows that exposure to more gender egalitarian policies affects labour market outcomes, division of domestic work and attitudes. This suggests that other gender equality promoting policies, such as more equal parental leave taking or further day care expansion to promote maternal employment, could have beneficial effects on both directly observed outcomes and gender norms in the long-run.

A potential issue of this chapter may seem that the results are specific to the unusual German context which has provided a rich field for researchers but with limited external validity. While the setting is unquestionably unique, the headline finding—that exposure to a gender-egalitarian or -traditional culture has strong impacts on child penalties—can arguably be generalised. A natural addition to this work would be using an epidemiological approach to study child penalties for different groups of immigrants within one country.

This chapter shows the impact of culture on child penalties, but it is unable to disentangle which aspects of socialisation are the strongest determinants of the magnitude of child penalties. This could either be parental norms which are passed on to their children, exposure to norms through peers or a long-run consequence of observing more gender egalitarian outcomes in everyday life. While socialisation can hardly be randomised, developing a better understanding of which factors impact child penalties and thus gender inequality as a whole, is a very promising avenue for future research.

Bibliography

- ABADIE, A., A. DIAMOND, AND J. HAINMUELLER (2010): “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program,” *Journal of the American Statistical Association*, 105, 493–505.
- ADDA, J., C. DUSTMANN, AND K. STEVENS (2017): “The Career Costs of Children,” *Journal of Political Economy*, 125, 293–337.
- AIKEN, L. S. AND S. G. WEST (1991): *Multiple regression: Testing and interpreting interactions*, Sage Publications.
- AKSOY, C. G., C. S. CARPENTER, R. DE HAAS, AND K. D. TRAN (2020): “Do laws shape attitudes? Evidence from same-sex relationship recognition policies in Europe,” *European Economic Review*, 124, 103399.
- ALBERINI, A. AND J. KAHN (2006): *Handbook on Contingent Valuation*, Edward Elgar Publishing.
- ALESINA, A. AND N. FUCHS-SCHÜNDELN (2007): “Good-bye Lenin (or Not?): The Effect of Communism on People’s Preferences,” *American Economic Review*, 97, 1507–1528.
- ALT, C., A. BETHMANN, B. GEDON, S. HUBERT, K. HÜSKEN, AND K. LIPPERT (2018): “Kinderbetreuungsstudie. Längsschnittdatensatz 2012-2016. Version: 1. DJI - Deutsches Jugendinstitut. Dataset.” .
- ALT, C., B. GEDON, S. HUBERT, K. HÜSKEN, AND K. LIPPERT (2019): *DJI-Kinderbetreuungsreport 2018.*, Deutsches Jugendinstitut (DJI), München.
- ALTONJI, J. G. AND R. M. BLANK (1999): “Race and gender in the labor market,” *Handbook of Labor Economics*, 3, 3143–3259.
- ANDERSEN, S. C. AND T. S. GUUL (2019): “Reducing Minority Discrimination at the Front Line—Combined Survey and Field Experimental Evidence,” *Journal of Public Administration Research and Theory*, 29, 429–444.

- ANGELOV, N., P. JOHANSSON, AND E. LINDAHL (2016): “Parenthood and the gender gap in pay,” *Journal of Labor Economics*, 34, 545–579.
- ARROW, K. J. (1973): *The Theory of Discrimination*, Princeton University Press.
- AUSPURG, K., T. HINZ, AND L. SCHMID (2017): “Contexts and conditions of ethnic discrimination: Evidence from a field experiment in a German housing market,” *Journal of Housing Economics*, 35, 26–36.
- AZOULAY, P., J. S. GRAFF ZIVIN, AND J. WANG (2010): “Superstar extinction,” *The Quarterly Journal of Economics*, 125, 549–589.
- BAKER, M., J. GRUBER, AND K. MILLIGAN (2008): “Universal child care, maternal labor supply, and family well-being,” *Journal of Political Economy*, 116, 709–745.
- (2019): “The Long-Run Impacts of a Universal Child Care Program,” *American Economic Journal: Economic Policy*, 11, 1–26.
- BAKER, M. AND K. MILLIGAN (2016): “Boy-girl differences in parental time investments: Evidence from three countries,” *Journal of Human Capital*, 10, 399–441.
- BANA, S. H., K. BEDARD, AND M. ROSSIN-SLATER (2020): “The Impacts of Paid Family Leave Benefits: Regression Kink Evidence from California Administrative Data,” *Journal of Policy Analysis and Management*, 39, 888–929.
- BARTEL, A. P., M. ROSSIN-SLATER, C. J. RUHM, M. SLOPEN, AND J. WALDFOGEL (2021): “The Impact of Paid Family Leave on Employers: Evidence from New York,” *IZA Discussion Paper*, 14262.
- BARTH, D., J. JESSEN, C. K. SPIESS, AND K. WROHLICH (2020): “Mothers in Eastern and Western Germany: Employment Rates and Attitudes Are Converging, Full-Time Employment is Not,” *DIW Weekly Report*, 10, 403–410.
- BASTIAN, J. AND L. LOCHNER (forthcoming): “The EITC and Maternal Time Use: More Time Working and Less Time with Kids?” *Journal of Labor Economics*.
- BATSAIKHAN, M., M. GØRTZ, J. KENNES, R. S. LYG, D. MONTE, AND N. TUMEN-NASAN (2019): “Daycare Choice and Ethnic Diversity: Evidence from a Randomized Survey,” *Aarhus University Economics Working Papers*, 2019-02.
- BAUCHMÜLLER, R., M. GØRTZ, AND A. W. RASMUSSEN (2014): “Long-run benefits from universal high-quality preschooling,” *Early Childhood Research Quarterly*, 29, 457–470.

- BAUERNSCHUSTER, S., T. HENER, AND H. RAINER (2016): “Children of a (policy) revolution: The introduction of universal child care and its effect on fertility,” *Journal of the European Economic Association*, 14, 975–1005.
- BAUERNSCHUSTER, S. AND H. RAINER (2012): “Political regimes and the family: how sex-role attitudes continue to differ in reunified Germany,” *Journal of Population Economics*, 25, 5–27.
- BAUERNSCHUSTER, S. AND M. SCHLOTTER (2015): “Public child care and mothers’ labor supply—Evidence from two quasi-experiments,” *Journal of Public Economics*, 123, 1–16.
- BAUM, C. L. AND C. J. RUHM (2016): “The effects of paid family leave in California on labor market outcomes,” *Journal of Policy Analysis and Management*, 35, 333–356.
- BAXTER, J. AND M. BITTMAN (1995): “Measuring time spent on housework: A comparison of two approaches,” *Australian Journal of Social Research*, 1, 21–46.
- BEBLO, M. AND L. GÖRGES (2018): “On the nature of nurture. The malleability of gender differences in work preferences,” *Journal of Economic Behavior & Organization*, 151, 19–41.
- BECKER, G. S. (1957): *The Economics of Discrimination*, University of Chicago press.
- (1981): *A Treatise on the Family*, Harvard University Press.
- BECKER, G. S., W. H. HUBBARD, AND K. M. MURPHY (2010): “Explaining the worldwide boom in higher education of women,” *Journal of Human Capital*, 4, 203–241.
- BECKER, S. O., L. MERGELE, AND L. WOESSMANN (2020): “The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism,” *Journal of Economic Perspectives*, 34, 143–171.
- BENNEDSEN, M., F. PÉREZ-GONZÁLEZ, AND D. WOLFENZON (2020): “Do CEOs matter? Evidence from hospitalization events,” *The Journal of Finance*, 75, 1877–1911.
- BERKES, J., J. JESSEN, AND F. WEINHARDT (2021): “When equality doesn’t make it indoors: Gender differences in market and non-market work in the GDR and reunified Germany,” *mimeo*.

- BERLINSKI, S. AND S. GALIANI (2007): “The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment,” *Labour Economics*, 14, 665–680.
- BERTRAND, M. (2020): “Gender in the Twenty-First Century,” *AEA Papers and Proceedings*, 110, 1–24.
- BERTRAND, M. AND E. DUFLO (2017): “Field Experiments on Discrimination,” *Handbook of Economic Field Experiments*, 1, 309–393.
- BERTRAND, M., C. GOLDIN, AND L. F. KATZ (2010): “Dynamics of the gender gap for young professionals in the financial and corporate sectors,” *American Economic Journal: Applied Economics*, 2, 228–55.
- BERTRAND, M., E. KAMENICA, AND J. PAN (2015): “Gender identity and relative income within households,” *The Quarterly Journal of Economics*, 130, 571–614.
- BIEWEN, M. AND S. SEIFERT (2018): “Potential Parenthood and Career Progression of Men and Women – A Simultaneous Hazards Approach,” *The B.E. Journal of Economic Analysis & Policy*, 18, 20170154.
- BISIN, A. AND T. VERDIER (2001): “The economics of cultural transmission and the dynamics of preferences,” *Journal of Economic Theory*, 97, 298–319.
- BJÖRKLUND, A. AND K. G. SALVANES (2011): “Education and family background: Mechanisms and policies,” in *Handbook of the Economics of Education*, Elsevier, vol. 3, 201–247.
- BLANDEN, J., E. DEL BONO, S. MCNALLY, AND B. RABE (2016): “Universal preschool education: The case of public funding with private provision,” *The Economic Journal*, 126, 682–723.
- BLAU, F. D. AND L. M. KAHN (2003): “Understanding international differences in the gender pay gap,” *Journal of Labor Economics*, 21, 106–144.
- (2013): “Female labor supply: Why is the United States falling behind?” *American Economic Review*, 103, 251–56.
- BLINDER, A. S. (1973): “Wage discrimination: reduced form and structural estimates,” *Journal of Human Resources*, 436–455.
- BLOSSFELD, H.-P. (1987): “Labor-market entry and the sexual segregation of careers in the Federal Republic of Germany,” *American Journal of Sociology*, 93, 89–118.

- BLOSSFELD, H.-P. AND J. VON MAURICE (2011): “Education as a lifelong process,” *Zeitschrift für Erziehungswissenschaft*, 14, 19–34.
- BMFSFJ (2008): “Elterngeld als Teil nachhaltiger Familienpolitik. Materialien aus dem Kompetenzzentrum für familienbezogene Leistungen im Bundesministerium für Familie, Senioren, Frauen und Jugend,” .
- (2018): “Family report 2017. Benefits, effects, trends.” .
- BOELMANN, B., A. RAUTE, AND U. SCHÖNBERG (2021): “Wind of Change? Cultural Determinants of Maternal Labor Supply,” *CREAM Discussion Paper 20/20*.
- BONEFELD, M. AND O. DICKHÄUSER (2018): “(Biased) grading of students’ performance: Students’ names, performance level, and implicit attitudes,” *Frontiers in Psychology*, 9, 481.
- BONEVA, T. AND C. RAUH (2018): “Parental Beliefs about Returns to Educational Investments—The Later the Better?” *Journal of the European Economic Association*, 16, 1669–1711.
- BOOTH, C. L., K. A. CLARKE-STEWART, D. L. VANDELL, K. MCCARTNEY, AND M. T. OWEN (2002): “Child-Care Usage and Mother-Infant “Quality Time”,” *Journal of Marriage and Family*, 64, 16–26.
- BORRA, C., M. BROWNING, AND A. SEVILLA (2021): “Marriage and housework,” *Oxford Economic Papers*, 73, 479–508.
- BRADBURY, B., M. CORAK, J. WALDFOGEL, AND E. WASHBROOK (2015): *Too many children left behind: The US achievement gap in comparative perspective*, Russell Sage Foundation.
- BRADLEY, R. H., R. F. CORWYN, M. BURCHINAL, H. P. MCADOO, AND C. G. COLL (2001): “The home environments of children in the United States Part II: Relations with behavioral development through age thirteen,” *Child Development*, 72, 1868–1886.
- BRENØE, A. A., S. P. CANAAN, N. A. HARMON, AND H. N. ROYER (2020): “Is Parental Leave Costly For Firms and Coworkers?” *NBER Working Paper Series*, 26622.
- BRÜDERL, J., S. DROBNIC, K. HANK, F. J. NEYER, S. WALPER, P. ALT, C. BOZOYAN, C. FINN, R. FRISTER, M. GARRET, T. GONZALEZ AVILES, H. GREISCHL, N. GRÖPLER, K. HAJEK, M. HERZIG, B. HUYER-MAY, R. LENKE,

- L. MINKUS, T. PETER, J. REIM, C. SCHMIEDEBERG, P. SCHÜTZE, N. SCHUMANN, C. THÖNISSEN, M. WETZEL, AND B. WILHELM (2020): “The German Family Panel (pairfam). GESIS Data Archive, Cologne. ZA5678 Data file Version 11.0.0,” .
- BUNDESVERFASSUNGSGERICHT (2003a): “1 BvR 302/96,” .
- (2003b): “Press release 98 / 2003-12-09,” .
- BURDA, M. C. AND S. SEELE (2017): “Das deutsche Arbeitsmarktwunder: Eine Bilanz,” *Perspektiven der Wirtschaftspolitik*, 18, 179–204.
- CAMEHL, G. F., J. F. STAHL, P. S. SCHÖBER, AND C. K. SPIESS (2015): “Does better, cheaper day care make for more satisfied parents?” *DIW Economic Bulletin*, 5, 604–611.
- CAMPA, P. AND M. SERAFINELLI (2019): “Politico-economic regimes and attitudes: Female workers under state socialism,” *Review of Economics and Statistics*, 101, 233–248.
- CARD, D., A. R. CARDOSO, AND P. KLINE (2015): “Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women,” *The Quarterly Journal of Economics*, 131, 633–686.
- CARNEIRO, P. M. AND J. J. HECKMAN (2003): “Human capital policy,” *NBER Working Paper Series*, 9495.
- CASCIO, E. U. (2009): “Maternal labor supply and the introduction of kindergartens into American public schools,” *Journal of Human Resources*, 44, 140–170.
- CATTANEO, M. D., M. JANSSON, AND X. MA (2018): “Manipulation testing based on density discontinuity,” *The Stata Journal*, 18, 234–261.
- CLARKE-STEWART, A., C. P. GRUBER, AND L. M. FITZGERALD (1994): *Children at home and in day care*, Psychology Press.
- COLLISCHON, M., A. EBERL, AND M. REICHEL (2020): “Structural legacies and the motherhood penalty: How past societal contexts shape mothers’ employment outcomes in reunified Germany,” *mimeo*.
- COLTRANE, S. (2000): “Research on household labor: Modeling and measuring the social embeddedness of routine family work,” *Journal of Marriage and Family*, 62, 1208–1233.

- COOKE, L. P. (2007): “Persistent policy effects on the division of domestic tasks in reunified Germany,” *Journal of Marriage and Family*, 69, 930–950.
- CORNELISSEN, T., C. DUSTMANN, A. RAUTE, AND U. SCHÖNBERG (2018): “Who benefits from universal child care? Estimating marginal returns to early child care attendance,” *Journal of Political Economy*, 126, 2356–2409.
- CORNELISSEN, T., C. DUSTMANN, AND U. SCHÖNBERG (2017): “Peer effects in the workplace,” *American Economic Review*, 107, 425–56.
- CORRIGALL, E. A. AND A. M. KONRAD (2007): “Gender role attitudes and careers: A longitudinal study,” *Sex Roles*, 56, 847–855.
- CÓRTEZ, P. AND J. PAN (2020): “Children and the Remaining Gender Gaps in the Labor Market,” *IZA Discussion Paper*, 13759.
- CRAIG, L. AND A. POWELL (2013): “Non-parental childcare, time pressure and the gendered division of paid work, domestic work and parental childcare,” *Community, Work & Family*, 16, 100–119.
- CUNHA, F., I. ELO, AND J. CULHANE (2013): “Eliciting maternal expectations about the technology of cognitive skill formation,” *NBER Working Paper series*, 19144.
- CUNHA, F., J. J. HECKMAN, L. LOCHNER, AND D. V. MASTEROV (2006): “Interpreting the evidence on life cycle skill formation,” *Handbook of the Economics of Education*, 1, 697–812.
- CUNNINGHAM, M. (2001): “Parental influences on the gendered division of housework,” *American Sociological Review*, 66, 184–203.
- DAHL, G. B., K. V. LØKEN, AND M. MOGSTAD (2014): “Peer effects in program participation,” *American Economic Review*, 104, 2049–74.
- DAHL, G. B., K. V. LØKEN, M. MOGSTAD, AND K. V. SALVANES (2016): “What is the case for paid maternity leave?” *Review of Economics and Statistics*, 98, 655–670.
- DATTA GUPTA, N. AND M. SIMONSEN (2010): “Non-cognitive child outcomes and universal high quality child care,” *Journal of Public Economics*, 94, 30–43.
- (2012): “The effects of type of non-parental child care on pre-teen skills and risky behavior,” *Economics Letters*, 116, 622–625.

- DEL BONO, E., M. FRANCESCONI, Y. KELLY, AND A. SACKER (2016): “Early maternal time investment and early child outcomes,” *The Economic Journal*, 126, F96–F135.
- DESTATIS (2008): “Öffentliche Sozialleistungen. Statistik zum Elterngeld. Elterngeld für Geburten 2007 Anträge von Januar 2007 bis Juni 2008.” *Statistisches Bundesamt*, Wiesbaden.
- DESTATIS (2015): “Zeitverwendungserhebung, Aktivitäten in Stunden und Minuten für ausgewählte Personengruppen 2012/2013,” *Statistisches Bundesamt*, Wiesbaden.
- DESTATIS (2017): “Kinder und tätige Personen in Tageseinrichtungen und öffentlich geförderter Kindertagespflege, verschiedene Jahre,” *Statistisches Bundesamt*, Wiesbaden.
- (2018): “Unternehmensstrukturstatistiken und Statistik für kleine und mittlere Unternehmen nach dem EU-Unternehmensbegriff.” *Statistisches Bundesamt*, Wiesbaden.
- (2021): “Elterngeld 2020: Väteranteil steigt auf knapp 25%,” *Statistisches Bundesamt*, Wiesbaden.
- DOBKIN, C., A. FINKELSTEIN, R. KLUENDER, AND M. J. NOTOWIDIGDO (2018): “The economic consequences of hospital admissions,” *American Economic Review*, 108, 308–52.
- DOEPKE, M., G. SORRENTI, AND F. ZILIBOTTI (2019): “The economics of parenting,” *Annual Review of Economics*, 11, 55–84.
- DOEPKE, M. AND F. ZILIBOTTI (2017): “Parenting with style: Altruism and paternalism in intergenerational preference transmission,” *Econometrica*, 85, 1331–1371.
- DOTTI SANI, G. M. AND J. TREAS (2016): “Educational gradients in parents’ child-care time across countries, 1965–2012,” *Journal of Marriage and Family*, 78, 1083–1096.
- DOYLE, O., C. HARMON, J. J. HECKMAN, C. LOGUE, AND S. H. MOON (2017): “Early skill formation and the efficiency of parental investment: A randomized controlled trial of home visiting,” *Labour Economics*, 45, 40–58.
- DRANGE, N. AND T. HAVNES (2019): “Early childcare and cognitive development: Evidence from an assignment lottery,” *Journal of Labor Economics*, 37, 581–620.

- DRANGE, N. AND K. TELLE (2020): “Segregation in a Universal Child Care System: Descriptive Findings from Norway,” *European Sociological Review*.
- ELANGO, S., J. L. GARCÍA, J. J. HECKMAN, AND A. HOJMAN (2015): “Early childhood education,” in *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, University of Chicago Press, 235–297.
- FANG, H. AND A. MORO (2011): “Theories of statistical discrimination and affirmative action: A survey,” *Handbook of Social Economics*, 1, 133–200.
- FARRÉ, L. AND F. VELLA (2013): “The intergenerational transmission of gender role attitudes and its implications for female labour force participation,” *Economica*, 80, 219–247.
- FARRÉ, L. AND L. GONZÁLEZ (2019): “Does paternity leave reduce fertility?” *Journal of Public Economics*, 172, 52–66.
- FELFE, C. AND R. LALIVE (2018): “Does early child care affect children’s development?” *Journal of Public Economics*, 159, 33–53.
- FELFE, C., N. NOLLENBERGER, AND N. RODRÍGUEZ-PLANAS (2015): “Can’t buy mommy’s love? Universal childcare and children’s long-term cognitive development,” *Journal of Population Economics*, 28, 393–422.
- FELFE, C. AND L. ZIEROW (2018): “From dawn till dusk: Implications of full-day care for children’s development,” *Labour Economics*, 55, 259–281.
- FERNÁNDEZ, R. (2007): “Women, work, and culture,” *Journal of the European Economic Association*, 5, 305–332.
- (2011): “Does culture matter?” *Handbook of Social Economics*, 1, 481–510.
- (2013): “Cultural change as learning: The evolution of female labor force participation over a century,” *American Economic Review*, 103, 472–500.
- FERNÁNDEZ, R. AND A. FOGLI (2009): “Culture: An empirical investigation of beliefs, work, and fertility,” *American Economic Journal: Macroeconomics*, 1, 146–77.
- FERNÁNDEZ, R., A. FOGLI, AND C. OLIVETTI (2004): “Mothers and sons: Preference formation and female labor force dynamics,” *The Quarterly Journal of Economics*, 119, 1249–1299.

- FERNÁNDEZ-KRANZ, D. AND N. RODRÍGUEZ-PLANAS (2021): “Too family friendly? The consequences of parent part-time working rights,” *Journal of Public Economics*, 197, 104407.
- FIEBIGER, H. (1991): “Die Zeitbudgeterhebung 1990 in der ehemaligen DDR,” *M. Ehling/Rv Schweitzer, Zeitbudgeterhebung der amtlichen Statistik*, 12–38.
- FIORINI, M. AND M. P. KEANE (2014): “How the allocation of children’s time affects cognitive and noncognitive development,” *Journal of Labor Economics*, 32, 787–836.
- FOGLI, A. AND L. VELDKAMP (2011): “Nature or nurture? Learning and the geography of female labor force participation,” *Econometrica*, 79, 1103–1138.
- FOLBRE, N. AND M. BITTMAN (2004): *Family time: The social organization of care*, vol. 2, Psychology Press.
- FORT, M., A. ICHINO, AND G. ZANELLA (2020): “Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families,” *Journal of Political Economy*, 128, 158–205.
- FORTIN, N., T. LEMIEUX, AND S. FIRPO (2011): “Decomposition methods in economics,” in *Handbook of Labor Economics*, Elsevier, vol. 4, 1–102.
- FORTIN, N. M. (2005): “Gender role attitudes and the labour-market outcomes of women across OECD countries,” *Oxford Review of Economic Policy*, 21, 416–438.
- FRODERMANN, C., K. WROHLICH, AND A. ZUCCO (2020): “Parental Leave Reform and Long-Run Earnings of Mothers,” *IZA Discussion Papers*, 12935.
- FUCHS-SCHÜNDELN, N. AND M. SCHÜNDELN (2005): “Precautionary savings and Self-Selection: Evidence from the German reunification “Experiment”,” *The Quarterly Journal of Economics*, 120, 1085–1120.
- (2009): “Who stays, who goes, who returns? East–West migration within Germany since reunification,” *Economics of Transition*, 17, 703–738.
- GALLEN, Y. (2019): “The effect of parental leave extensions on firms and coworkers,” *Working Paper*.
- GANGL, M. AND A. ZIEFLE (2009): “Motherhood, labor force behavior, and women’s careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States,” *Demography*, 46, 341–369.

- GELBACH, J. B. (2002): “Public schooling for young children and maternal labor supply,” *American Economic Review*, 92, 307–322.
- (2016): “When do covariates matter? And which ones, and how much?” *Journal of Labor Economics*, 34, 509–543.
- GERSHUNY, J., M. VEGA-RAPUN, AND J. LAMOTE (2020): *Multinational Time Use Study*, Centre for Time Use Research, UCL IOE, University College London.
- GIMENEZ-NADAL, J. I. AND J. MOLINA (2020): “The Gender Gap in Time Allocation in Europe,” *IZA Discussion Paper*, 13461.
- GIMENEZ-NADAL, J. I. AND J. A. MOLINA (2013): “Parents’ education as a determinant of educational childcare time,” *Journal of Population Economics*, 26, 719–749.
- GIMENEZ-NADAL, J. I. AND A. SEVILLA (2012): “Trends in time allocation: A cross-country analysis,” *European Economic Review*, 56, 1338–1359.
- GINJA, R., J. JANS, AND A. KARIMI (2020a): “Parental leave benefits, household labor supply, and children’s long-run outcomes,” *Journal of Labor Economics*, 38, 261–320.
- GINJA, R., A. KARIMI, AND P. XIAO (2020b): “Employer Responses to Family Leave Programs,” *Working Paper*.
- GIULIANO, P. (2021): “Gender and culture,” *Oxford Review of Economic Policy*, 36, 944–961.
- GOEBEL, J., M. M. GRABKA, S. LIEBIG, M. KROH, D. RICHTER, C. SCHRÖDER, AND J. SCHUPP (2019): “The German Socio-Economic Panel (SOEP),” *Jahrbücher für Nationalökonomie und Statistik*, 239, 345–360.
- GOLDIN, C. (2014): “A grand gender convergence: Its last chapter,” *American Economic Review*, 104, 1091–1119.
- GRUNOW, D., K. BEGALL, AND S. BUCHLER (2018): “Gender ideologies in Europe: A multidimensional framework,” *Journal of Marriage and Family*, 80, 42–60.
- GUIISO, L., P. SAPIENZA, AND L. ZINGALES (2006): “Does culture affect economic outcomes?” *Journal of Economic Perspectives*, 20, 23–48.
- GURRYAN, J., E. HURST, AND M. KEARNEY (2008): “Parental education and parental time with children,” *Journal of Economic Perspectives*, 22, 23–46.

- HABIBOV, N. AND J. COYLE (2014): “Effect of early child care and education attendance hours on parent–child interaction in Central Asia: Evaluation of multiple indicator cluster surveys,” *Child & Youth Services*, 35, 169–191.
- HAKIM, C. (2000): *Work-lifestyle choices in the 21st century: Preference theory*, OUP Oxford.
- HANK, K., M. KREYENFELD, AND C. K. SPIESS (2004): “Kinderbetreuung und Fertilität in Deutschland/Child Care and Fertility in Germany,” *Zeitschrift für Soziologie*, 33, 228–244.
- HAVNES, T. AND M. MOGSTAD (2011a): “Money for nothing? Universal child care and maternal employment,” *Journal of Public Economics*, 95, 1455–1465.
- (2011b): “No child left behind: Subsidized child care and children’s long-run outcomes,” *American Economic Journal: Economic Policy*, 3, 97–129.
- (2015): “Is universal child care leveling the playing field?” *Journal of Public Economics*, 127, 100–114.
- HEINING, J., W. KLOSTERHUBER, P. LEHNERT, S. SETH, ET AL. (2016): “Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell 1993-2014 (LIAB LM 9314),” Tech. rep., Institute for Employment Research (IAB), Nuremberg, Germany.
- HEINZE, A. AND E. WOLF (2010): “The intra-firm gender wage gap: A new view on wage differentials based on linked employer–employee data,” *Journal of Population Economics*, 23, 851–879.
- HEISIG, K. AND L. ZIEROW (2021): “Parental Leave and Long-Term Life Satisfaction of Children - Quasi-Experimental Evidence from Former East Germany,” *CESifo Working Paper*, 7806.
- HENSVIK, L. AND O. ROSENQVIST (2019): “Keeping the Production Line Running - Internal Substitution and Employee Absence,” *Journal of Human Resources*, 54, 200–224.
- HERBST, C. M. AND E. TEKIN (2014): “Child care subsidies, maternal health, and child-parent interactions: Evidence from three nationally representative datasets,” *Health Economics*, 23, 894–916.
- HERMES, H., P. LERGETPORER, F. PETER, AND S. WIEDERHOLD (2021): “Behavioral Barriers and the Socioeconomic Gap in Child Care Enrollment,” *mimeo*.

- HERZOG, S. AND T. KLEIN (2018): “Matching practices for childcare - Germany,” *Matching in practice, MiP Country Profile 26*.
- HIRSCH, B. AND S. MUELLER (2012): “The productivity effect of temporary agency work: Evidence from German panel data,” *The Economic Journal*, 122, F216–F235.
- HUEBENER, M., J. JESSEN, D. KUEHNLE, AND M. OBERFICHTNER (2021): “A Firm-Side Perspective on Parental Leave,” *IZA Discussion Paper*, 14478.
- HUEBENER, M., D. KUEHNLE, AND C. K. SPIESS (2019): “Parental leave policies and socio-economic gaps in child development: Evidence from a substantial benefit reform using administrative data,” *Labour Economics*, 61, 101754.
- HUEBENER, M., A. PAPE, AND C. K. SPIESS (2020): “Parental labour supply responses to the abolition of day care fees,” *Journal of Economic Behavior & Organization*, 180, 510–543.
- HUININK, J., J. BRÜDERL, B. NAUCK, S. WALPER, L. CASTIGLIONI, AND M. FELDHAUS (2011): “Panel analysis of intimate relationships and family dynamics (pairfam): Conceptual framework and design,” *ZfF–Zeitschrift für Familienforschung/Journal of Family Research*, 23.
- JACOBEBBINGHAUS, P. AND S. SETH (2007): “The German integrated employment biographies sample IEBS,” *Schmollers Jahrbuch*, 127, 335–342.
- JÄGER, S. AND J. HEINING (2019): “How Substitutable are Workers? Evidence from Worker Deaths,” *Working Paper*.
- JANN, B. (2008): “A Stata implementation of the Blinder-Oaxaca decomposition,” *Stata Journal*, 8, 453–479.
- JARAVEL, X., N. PETKOVA, AND A. BELL (2018): “Team-specific capital and innovation,” *American Economic Review*, 108, 1034–73.
- JESSEN, J. (2021): “Culture, Children and Couple Gender Inequality,” *DIW Berlin Discussion Paper*, 1957.
- JESSEN, J., R. JESSEN, AND J. KLUVE (2019): “Punishing potential mothers? Evidence for statistical employer discrimination from a natural experiment,” *Labour Economics*, 59, 164–172.
- JESSEN, J., S. SCHMITZ, C. K. SPIESS, AND S. WAIGHTS (2018): “Kita-Besuch hängt trotz ausgeweitetem Rechtsanspruch noch immer vom Familienhintergrund ab,” *DIW-Wochenbericht*, 85, 825–835.

- JESSEN, J., S. SCHMITZ, AND S. WAIGHTS (2020a): “Understanding day care enrolment gaps,” *Journal of Public Economics*, 190, 104252.
- JESSEN, J., C. K. SPIESS, AND S. WAIGHTS (2020b): “Center-Based Care and Parenting Activities,” *DIW Berlin Discussion Paper*, 1897.
- JÜRGES, H. (2017): “Financial incentives, timing of births, and infant health: a closer look into the delivery room,” *The European Journal of Health Economics*, 18, 195–208.
- KAAS, L. AND C. MANGER (2012): “Ethnic discrimination in Germany’s labour market: A field experiment,” *German Economic Review*, 13, 1–20.
- KALB, G. AND J. C. VAN OURS (2014): “Reading to young children: A head-start in life?” *Economics of Education Review*, 40, 1–24.
- KALIL, A., R. RYAN, AND M. COREY (2012): “Diverging destinies: Maternal education and the developmental gradient in time with children,” *Demography*, 49, 1361–1383.
- KAUL, A., S. KLÖSSNER, G. PFEIFER, AND M. SCHIELER (2021): “Standard Synthetic Control Methods: The Case Of Using All Pre-Intervention Outcomes Together With Covariates,” *Journal of Business & Economic Statistics*, 1–34.
- KIM, J. H., W. SCHULZ, T. ZIMMERMANN, AND K. HAHLEWEG (2018): “Parent–child interactions and child outcomes: Evidence from randomized intervention,” *Labour Economics*, 54, 152–171.
- KIMMEL, J. (1998): “Child care costs as a barrier to employment for single and married mothers,” *Review of Economics and Statistics*, 80, 287–299.
- KITTERØD, R. H., T. H. LYGSTAD, ET AL. (2005): “Diary versus questionnaire information on time spent on housework—The case of Norway,” *Electronic International Journal of Time Use Research*, 2, 13–32.
- KLEVEN, H. AND C. LANDAIS (2017): “Gender inequality and economic development: fertility, education and norms,” *Economica*, 84, 180–209.
- KLEVEN, H., C. LANDAIS, J. POSCH, A. STEINHAEUER, AND J. ZWEIMÜLLER (2019a): “Child penalties across countries: Evidence and explanations,” *AEA Papers and Proceedings*, 109, 122–26.

- (2020): “Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation,” *NBER Working Paper series*, 28082.
- KLEVEN, H., C. LANDAIS, AND J. E. SØGAARD (2018): “Children and gender inequality: Evidence from Denmark,” *NBER Working Paper series*, 24219.
- (2019b): “Children and gender inequality: Evidence from Denmark,” *American Economic Journal: Applied Economics*, 11, 181–209.
- KLINE, P. AND C. R. WALTERS (2016): “Evaluating public programs with close substitutes: The case of Head Start,” *The Quarterly Journal of Economics*, 131, 1795–1848.
- KLÜSENER, S. AND J. R. GOLDSTEIN (2016): “A long-standing demographic East–West divide in Germany,” *Population, Space and Place*, 22, 5–22.
- KLUBE, J. AND S. SCHMITZ (2018): “Back to work: Parental benefits and mothers’ labor market outcomes in the medium run,” *ILR Review*, 71, 143–173.
- KLUBE, J. AND M. TAMM (2013): “Parental leave regulations, mothers’ labor force attachment and fathers’ childcare involvement: Evidence from a natural experiment,” *Journal of Population Economics*, 26, 983–1005.
- KNOWLES, J. A. (2013): “Why are married men working so much? An aggregate analysis of intra-household bargaining and labour supply,” *The Review of Economic Studies*, 80, 1055–1085.
- KOOPMANS, R., S. VEIT, AND R. YEMANE (2019): “Taste or statistics? A correspondence study of ethnic, racial and religious labour market discrimination in Germany,” *Ethnic and Racial Studies*, 42, 233–252.
- KOSFELD, R. AND A. WERNER (2012): “Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011,” *Raumforschung und Raumordnung*, 70, 49–64.
- KRAPF, M., A. ROTH, AND M. SLOTWINSKI (2020): “The effect of childcare on parental earnings trajectories,” *CESifo Working Paper*, 8764.
- KRÖLL, A. AND R. BORCK (2013): “The Influence of Child Care on Maternal Health and Mother-Child Interaction,” *CESifo Working Paper*, 4289.
- KRUEGER, A. B. AND J.-S. PISCHKE (1995): “A comparative analysis of East and West German labor markets: before and after unification,” in *Differences and changes in wage structures*, University of Chicago Press, 405–446.

- KUEHNLE, D. AND M. OBERFICHTNER (2020): “Does Starting Universal Childcare Earlier Influence Children’s Skill Development?” *Demography*, 57, 1–38.
- KUGER, S., J. MARCUS, AND C. K. SPIESS (2019): “Day care quality and changes in the home learning environment of children,” *Education Economics*, 27, 265–286.
- KUZIEMKO, I., J. PAN, J. SHEN, AND E. WASHINGTON (2018): “The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?” *NBER Working Paper*, 24740.
- LALIVE, R. AND J. ZWEIMÜLLER (2009): “How does parental leave affect fertility and return to work? Evidence from two natural experiments,” *The Quarterly Journal of Economics*, 124, 1363–1402.
- LAUDENBACH, C., U. MALMENDIER, AND A. NIESSEN-RUENZI (2020): “The long-lasting effects of experiencing communism on attitudes towards financial markets,” *CEPR Discussion Paper No. DP14939*.
- LAZEAR, E. P. (2009): “Firm-Specific Human Capital: A Skill-Weights Approach,” *Journal of Political Economy*, 117, 914–940.
- LIPPMANN, Q., A. GEORGIEFF, AND C. SENIK (2020): “Undoing gender with institutions: Lessons from the German division and reunification,” *The Economic Journal*, 130, 1445–1470.
- LOEB, S., M. BRIDGES, D. BASSOK, B. FULLER, AND R. W. RUMBERGER (2007): “How much is too much? The influence of preschool centers on children’s social and cognitive development,” *Economics of Education Review*, 26, 52–66.
- LÜDEMANN, E. AND G. SCHWERDT (2013): “Migration background and educational tracking,” *Journal of Population Economics*, 26, 455–481.
- LUNDBERG, S., S. W. PABILONIA, AND J. WARD-BATTS (2007): “Time allocation of parents and investments in sons and daughters,” *Unpublished Manuscript*.
- LUNDBERG, S. AND R. A. POLLAK (1993): “Separate spheres bargaining and the marriage market,” *Journal of Political Economy*, 101, 988–1010.
- (1996): “Bargaining and distribution in marriage,” *Journal of Economic Perspectives*, 10, 139–158.
- LUNDBERG, S. AND E. ROSE (2000): “Parenthood and the earnings of married men and women,” *Labour Economics*, 7, 689–710.

- MAGNUSON, K. AND J. WALDFOGEL (2016): “Trends in income-related gaps in enrollment in early childhood education: 1968 to 2013,” *AERA Open*, 2, 2332858416648933.
- MAIER, L. (2014): “Methodik und Durchführung der Zeitverwendungserhebung 2012/2013,” *Wirtschaft und Statistik*, 11, 672–679.
- MANNING, A. (2011): “Imperfect competition in the labor market,” in *Handbook of Labor Economics*, Elsevier, vol. 4, 973–1041.
- MARINI, M. M. AND B. A. SHELTON (1993): “Measuring household work: Recent experience in the United States,” *Social Science Research*, 22, 361–382.
- MAYER, S. E., A. KALIL, P. OREOPOULOS, AND S. GALLEGOS (2019): “Using behavioral insights to increase parental engagement the parents and children together intervention,” *Journal of Human Resources*, 54, 900–925.
- MORONI, G., C. NICOLETTI, AND E. TOMINEY (2019): “Child Socio-Emotional Skills: The Role of Parental Inputs,” *IZA Discussion Paper*, 12432.
- MÜLLER, D. AND K. STRAUCH (2017): “Identifying mothers in administrative data,” *FDZ-Methodenreport*.
- MÜLLER, K.-U. AND K. WROHLICH (2020): “Does Subsidized Care for Toddlers Increase Maternal Labor Supply?: Evidence from a Large-Scale Expansion of Early Childcare,” *Labour Economics*, 62.
- NEUGART, M. AND H. OHLSSON (2013): “Economic incentives and the timing of births: Evidence from the German parental benefit reform 2007,” *Journal of Population Economics*, 26, 87–108.
- NICKEL, H. M. (1992): “"Mitgestalterinnen des Sozialismus"—Frauenarbeit in der DDR,” *Frauen in Deutschland*, 233–256.
- NICOLETTI, C. AND V. TONEI (2020): “Do parental time investments react to changes in child’s skills and health?” *European Economic Review*, 103491.
- NOLLENBERGER, N. AND N. RODRÍGUEZ-PLANAS (2015): “Full-time universal child-care in a context of low maternal employment: Quasi-experimental evidence from Spain,” *Labour Economics*, 36, 124–136.
- NYT (2021): “What’s in Biden’s Spending Plan: Free Preschool and National Paid Leave,” Issue: 2021-04-28.

- OAXACA, R. (1973): “Male-female wage differentials in urban labor markets,” *International Economic Review*, 14, 693–709.
- OECD (2016a): “OECD family database: Enrolment in childcare and pre-school,” *OECD - Social Policy Division - Directorate of Employment, Labour and Social Affairs*.
- (2016b): “OECD family database: Maternal employment,” *OECD - Social Policy Division - Directorate of Employment, Labour and Social Affairs*.
- (2017): *Starting Strong 2017-Key OECD Indicators on Early Childhood Education and Care*, OECD Publishing, Paris.
- (2018): *Engaging Young Children: Lessons from Research about Quality in Early Childhood Education and Care*, OECD Publishing, Paris.
- (2019): “Gender wage gap (indicator).” doi: 10.1787/7cee77aa-en, accessed on 21 July, 2021.
- (2020): “Net childcare costs (indicator),” doi: 10.1787/e328a9ee-en, accessed on 21 July, 2021.
- OLIVETTI, C. AND B. PETRONGOLO (2008): “Unequal pay or unequal employment? A cross-country analysis of gender gaps,” *Journal of Labor Economics*, 26, 621–654.
- (2016): “The evolution of gender gaps in industrialized countries,” *Annual review of Economics*, 8, 405–434.
- (2017): “The economic consequences of family policies: lessons from a century of legislation in high-income countries,” *Journal of Economic Perspectives*, 31, 205–30.
- ONDRICH, J., C. K. SPIESS, AND Q. YANG (1996): “Barefoot and in a German kitchen: Federal parental leave and benefit policy and the return to work after childbirth in Germany,” *Journal of Population Economics*, 9, 247–266.
- OSTER, E. (2019): “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 37, 187–204.
- PHELPS, E. S. (1972): “The statistical theory of racism and sexism,” *American Economic Review*, 62, 659–661.
- POLLAK, R. A. (2005): “Bargaining Power in Marriage: Earnings, Wage Rates and Household Production,” *NBER Working Papers series*, 11239.

- PRICE, J. AND A. KALIL (2019): “The effect of mother–child reading time on children’s reading skills: Evidence from natural within-family variation,” *Child Development*, 90, e688–e702.
- PRILLER, E. (1993): “Zeitverwendung im Wandel: Analysen mit Zeitbudgetdaten der DDR-Statistik,” in *Amtliche Daten der DDR und der neuen Bundesländer: Informationsquelle für die Sozialwissenschaften*, Mannheim: ZUMA, 46–64.
- RAUTE, A. (2019): “Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits,” *Journal of Public Economics*, 169, 203–222.
- RINDFUSS, R. R., D. K. GUILKEY, S. P. MORGAN, AND Ø. KRAVDAL (2010): “Child-care availability and fertility in Norway,” *Population and Development Review*, 36, 725–748.
- ROSENFELD, R. A., H. TRAPPE, AND J. C. GORNICK (2004): “Gender and work in Germany: Before and after reunification,” *Annual Review of Sociology*, 30, 103–124.
- ROSSIN-SLATER, M. (2018): “Maternity and Family Leave Policy,” in *The Oxford Handbook of Women and the Economy*, ed. by S. L. Averett, L. M. Argys, and S. D. Hoffman, New York: Oxford University Press, 1–24.
- SANDNER, M. AND S. L. THOMSEN (2020): “Preventing Child Maltreatment: Beneficial Side Effects of Public Childcare Provision,” *Discussion paper - Hannover Economics Papers (HEP)*, 1–45.
- SCHMIDHEINY, K. AND S. SIEGLOCH (2020): “On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization,” *ZEW-Centre for European Economic Research Discussion Paper*.
- SCHMITZ, S. (2020): “The Impact of Publicly Funded Childcare on Parental Well-Being: Evidence from Cut-Off Rules,” *European Journal of Population*, 36, 171–196.
- SCHMITZ, S. AND C. K. SPIESS (2021): “The intergenerational transmission of gender norms—why and how adolescent males with working mothers matter for female labour market outcomes,” *Socio-Economic Review*, 1–39.
- SCHMITZ, S., C. K. SPIESS, AND J. F. STAHL (2017): “Day care centers: Family expenditures increased significantly at some points between 1996 and 2015,” *DIW Economic Bulletin*, 7, 411–423.
- SCHMITZ, S. AND F. WEINHARDT (2019): “Immigration and the Evolution of Local Cultural Norms,” *IZA DP No.*, 12509.

- SCHOBER, P. S. (2014): “Parental leave and domestic work of mothers and fathers: A longitudinal study of two reforms in West Germany,” *Journal of Social Policy*, 43, 351.
- SCHOBER, P. S. AND C. K. SPIESS (2013): “Early childhood education activities and care arrangements of disadvantaged children in Germany,” *Child Indicators Research*, 6, 709–735.
- SCHROEDER, C., C. K. SPIESS, AND J. STORCK (2015): “Private spending on children’s education: Low-income families pay relatively more,” *DIW Economic Bulletin*, 5, 113–123.
- SCHÖNBERG, U. AND J. LUDSTECK (2014): “Expansions in maternity leave coverage and mothers’ labor market outcomes after childbirth,” *Journal of Labor Economics*, 32, 469–505.
- SEVILLA, A. (2021): “Gender Economics: An Assessment,” *Oxford Review of Economic Policy*, 36, 725–742.
- SIMINSKI, P. AND R. YETSENGA (2020): “Rethinking Specialisation and the Sexual Division of Labour in the 21st Century,” *IZA Discussion Paper*, 12977.
- SØRENSEN, A. AND H. TRAPPE (1995): “The persistence of gender inequality in earnings in the German Democratic Republic,” *American Sociological Review*, 398–406.
- SPIESS, C. K. (2008): “Early childhood education and care in Germany: The status quo and reform proposals,” *Journal of Business Economics, ZfB Special*, 1–21.
- (2013): “Investments in education: The early years offer great potential,” *DIW Economic Bulletin*, 3, 3–10.
- SPIESS, C. K. AND K. WROHLICH (2005): “Wie viele Kinderbetreuungsplätze fehlen in Deutschland?: Neue Bedarfsermittlung für Kinder unter drei Jahren auf der Basis von Mikrodaten,” *DIW Wochenbericht*, 72, 223–227.
- SPRENGHOLZ, M., A. WIEBER, AND E. HOLST (2020): “Gender identity and wives’ labor market outcomes in West and East Germany between 1983 and 2016,” *Socio-Economic Review*.
- STAHL, J. F., P. S. SCHOBER, AND C. K. SPIESS (2018): “Parental socio-economic status and childcare quality: Early inequalities in educational opportunity?” *Early Childhood Research Quarterly*, 44, 304–317.

- TAMM, M. (2013): “The impact of a large parental leave benefit reform on the timing of birth around the day of implementation,” *Oxford Bulletin of Economics and Statistics*, 75, 585–601.
- (2019): “Fathers’ parental leave-taking, childcare involvement and labor market participation,” *Labour Economics*, 59, 184–197.
- 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*.
- TODD, P. E. AND K. I. WOLPIN (2007): “The production of cognitive achievement in children: Home, school, and racial test score gaps,” *Journal of Human capital*, 1, 91–136.
- TRAPPE, H. (1996): “Work and family in women’s lives in the German Democratic Republic,” *Work and Occupations*, 23, 354–377.
- VERMEULEN, F. (2002): “Collective household models: principles and main results,” *Journal of Economic Surveys*, 16, 533–564.
- WALDFOGEL, J. (1998): “Understanding the " family gap" in pay for women with children,” *Journal of Economic Perspectives*, 12, 137–156.
- WEICHSELBAUMER, D. (2020): “Multiple Discrimination against Female Immigrants Wearing Headscarves,” *ILR Review*, 73, 600–627.
- WEICHSELBAUMER, D. AND R. WINTER-EBMER (2005): “A meta-analysis of the international gender wage gap,” *Journal of Economic Surveys*, 19, 479–511.
- (2006): “Rhetoric in economic research: The case of gender wage differentials,” *Industrial Relations: A Journal of Economy and Society*, 45, 416–436.
- WEINERT, S., H.-G. ROSSBACH, G. FAUST, H.-P. BLOSSFELD, AND C. ARTELT (2013): “Educational Processes, Competence Development and Selection Decisions in Preschool and School Age (BiKS-3-10),” *Data Set*.
- WELTEKE, C. AND K. WROHLICH (2019): “Peer effects in parental leave decisions,” *Labour Economics*, 57, 146–163.
- WEST, C. AND D. H. ZIMMERMAN (1987): “Doing gender,” *Gender & society*, 1, 125–151.

- WROHLICH, K. (2008): “The excess demand for subsidized child care in Germany,” *Applied Economics*, 40, 1217–1228.
- ZOCH, G. (2021): “Thirty Years after the Fall of the Berlin Wall—Do East and West Germans Still Differ in Their Attitudes to Female Employment and the Division of Housework?” *European Sociological Review*, 1–20.
- ZOCH, G. AND I. HONDRALIS (2017): “The expansion of low-cost, state-subsidized childcare availability and mothers’ return-to-work behaviour in East and West Germany,” *European Sociological Review*, 33, 693–707.
- ZOCH, G. AND P. S. SCHOBBER (2018): “Public child-care expansion and changing gender ideologies of parents in Germany,” *Journal of Marriage and Family*, 80, 1020–1039.