DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 46, ARTICLE 24, PAGES 693–722
PUBLISHED 7 APRIL 2022
https://www.demographic-research.org/Volumes/Vol46/24/
DOI: 10.4054/DemRes.2022.46.24

Research Article

Where does public childcare boost female labor force participation? Exploring geographical heterogeneity across Germany 2007–2017

Franz Neuberger

Tobias Rüttenauer

Martin Bujard

© 2022 Franz Neuberger, Tobias Rüttenauer & Martin Bujard.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.
See https://creativecommons.org/licenses/by/3.0/de/legalcode.

Contents

Demographic Research: Volume 46, Article 24
Research Article

Where does public childcare boost female labor force participation? Exploring geographical heterogeneity across Germany 2007–2017

Franz Neuberger[1]

Tobias Rüttenauer[2]

Martin Bujard[3]

Abstract

BACKGROUND
Public childcare provision and female labor force participation (FLP) have strongly increased over the past decades in European societies. However, studies offer heterogeneous findings on the link between public childcare and FLP.

OBJECTIVE
We investigate the link between public childcare and FLP, using different indicators of childcare and accounting for heterogeneous time trends and regional heterogeneity.

METHODS
Based on a balanced panel of all German counties from 2007 to 2017, we estimate the effect of an increasing enrollment rate for children aged 0–2 and 3–5 on FLP. We compare fixed effect (FE) and fixed effect individual slope estimators (FEIS) to control for county-specific time trends. Subsequently, we compare the results across regions with different levels of urbanization.

RESULTS
We find that most FE results are biased due to selection on trends. Still, when accounting for selection on trends, childcare enrollment for the age group 0–2 increases FLP in West Germany and in urban areas. Furthermore, childcare enrollment for children aged 3–5 years is associated with higher FLP in West Germany, in rural and, most strongly, in metropolitan areas.

CONCLUSIONS
Our study highlights important heterogeneity in the general time trends of FLP and the effectiveness of childcare arrangements across different regions in Germany.

[1] German Youth Institute (DJI), Munich, Germany. Email: fneuberger@dji.de.
[2] Nuffield College, University of Oxford, UK.
[3] Federal Institute for Population Research (BiB), Wiesbaden, Germany.

CONTRIBUTION

We provide a potential explanation for mixed results in previous studies. Moreover, our findings demonstrate that the effectiveness of childcare arrangements varies with the level of urbanization. Taking this heterogeneity into account can help to develop target-oriented policy interventions.

1. Introduction

The provision of public childcare is an essential topic for family policies in many countries (Immervoll and Barber 2006). For decades, the percentage of children in day-care facilities has increased in all OECD countries, with Germany recording one of the steepest rises (OECD 2017). Several European countries have adopted a "child-centered social investment strategy" (Esping-Andersen 2002: 26), whereby public childcare is supposed to raise female labor force participation (FLP), increase fertility, and foster child development. FLP has increased in the last decades (Thévenon 2013). Both developments – increases in childcare policies and increased FLP – are strongly related: together with female education and egalitarian gender attitudes they can be interpreted as part of the gender revolution (England 2010), which has occurred at different paces both within and between countries. A crucial question, therefore, is to identify the strength of the relationship between childcare policies and FLP and its patterns regarding time and regional type within a country. This paper focuses on the effect of the rise in public childcare coverage on female employment rates, as proposed by the economic model of the family (e.g., Becker 1991; Blau 2003) and the structural model of welfare state changes (Esping-Andersen 2002, 2009).

The childcare reforms in Germany are well suited to analyzing the effects of childcare policies on FLP, for three reasons: (1) childcare places for toddlers (0–2) more than doubled during the analyzed decade (2007–2017) in West Germany, (2) the share of full-time childcare places for children aged 3–5 increased by 57% in this decade, and (3) Germany is a federal country with heterogeneous regions that have vastly different enrollment rates for childcare and rich data. Thus, a wide range of literature studies the extensive and intensive margins of public childcare in Germany (Boll and Lagemann 2019; Müller and Wrohlich 2020; Schober and Spieß 2015; Zoch 2020; Zoch and Hondralis 2017), including research on the moderating effects of policy reforms, costs, quality of care, availability, and the influence on gender ideologies (Zoch and Schober 2018). The studies usually utilize individual data, while some more recent studies utilize within-country variation in macro-data panels. These macro-data approaches (Boll and

Lagemann 2019: Müller and Wrohlich 2020) extend existing knowledge by showing the macro-level returns to changes in public childcare over time.

However, we find that three aspects are given little attention in the research. First, current research mostly relies on classical fixed effect (FE) models, which can produce biased estimates if not controlled for selection on heterogeneous trends. Second, this heterogeneity could result not only from the historical divide between East and West Germany but also from regional differences such as between rural and urban areas – the importance of which was recently emphasized by Daniel et al. (2019). Third, current research usually focuses on possible effects from childcare for children aged 0–2, and neglects the much more common kindergarten.

To close this research gap, we utilize the more advanced fixed effects individual slope (FEIS) models (Brüderl and Ludwig 2015; Rüttenauer and Ludwig 2020; Wooldridge 2015). We utilize a full balanced panel on county-level data in Germany over 11 years including information on the overall childcare coverage and share of full-time places (with a care time of 7 and more hours per day) for both toddlers (age group 0–2) and kindergarten (age group 3–5) enrollment, covering the entire sector of early childhood care. This allows estimating the within-county effect of changes across the entire early childcare system on the FLP rate based on almost every German county from 2007 to 2017 and enables us to consider all types of spatial differentiation, as our data covers the entire country.

We contribute to the literature in the following ways. First, we extend research on childcare and female employment by using official administrative data at the county level. Second, we utilize FEIS models to account for heterogeneous time trends in FLP and childcare provision between counties, thereby relaxing an important assumption that is often violated with conventional FE models. Third, we examine differences in childcare systems due to the historical divide between West and East Germany and also across counties with different degrees of urbanization. Fourth, we consider changes in childcare provision for children aged 3–5, which has been rather neglected, especially in discussions on Germany, as the political focus has been on the childcare expansion for toddlers.

2. Public childcare and female employment

2.1 Policy reforms and the development of public childcare

Traditionally, Germany is a country with low fertility and low labor force participation of mothers and is historically characterized by strong regional heterogeneity, which is usually explained by the historical divide between East and West Germany.

Consequently, much research also utilizes this difference. In former West Germany (Müller and Wrohlich 2020) social norms and tax systems favored a classical male breadwinner model for families, while in former East Germany women were more frequently employed and childcare was more available. To close these gaps, in a paradigm shift in family policy, public childcare was expanded, changes were made to family law, and in 2007 generous parental leave was introduced, which influenced mothers' work behavior and preferences (Gangl and Ziefle 2015). In 2004 the Daycare Expansion Law (Tagesbetreuungsausbaugesetz, TAG) explicitly addressed the expansion of day care for children under the age of 3, formulating quality standards for child-minders. In 2008 the Childcare Funding Act (Kinderförderungsgesetz, KiFöG) prescribed a gradual expansion of childcare for children under the age of 3, denominating 1 August 2013 as the deadline for when supply had to meet demand. Since August 2013 each child under the age of 3 has been legally entitled to have access to public childcare. These laws resulted in Germany having one of the steepest increases in public childcare among all OECD countries (OECD 2017).

Overall, German social policy has created favorable conditions for the reconciliation of family and work, leading to a significant relief for parents and an overall increase in the quality of life for German mothers and fathers (Preisner et. al. 2018, 2020). Moreover, less affluent households are exempt from the costs (Schmitz, Spieß, and Stahl 2017). Thus, fees do not play a major role in the decision to utilize childcare provision, and only 23% (age group 0–2) and 20% (age group 3–5) of non-users mention fees as a reason for their non-utilization (Alt et al. 2017). These achievements look impressive at first, but the vast majority of children aged 0 to 2 in West Germany and around half of the children of that age group in East Germany are still in informal care arrangements: at home with their mothers, being cared for by their grandparents, or in other private care arrangements. The situation is quite different for children aged 3–5: the majority use professional care arrangements, but there is still unmet demand for full day-care places (Alt et al. 2017).

2.2 Theory and empirical research on public childcare and female employment

Theoretically, we follow the microeconomic and structural welfare state approaches. According to the economic model of the family (e.g., Becker 1991; Blau 2003), decisions regarding mothers' employment are framed by maximizing the utility of time invested in employment and care (among other things), given time and budget constraints. For couples, the "division of labor between market and household activities" (Becker 1991: 3), the latter including childcare, is a crucial decision, often resulting in specialization to increase productivity. Historically, women more often specialize in caring activities. However, the growing female earning power challenges this division (Becker 1991:

355ff.) and increases the opportunity costs. If a substantial share of childcare can be provided by formal or informal arrangements, both mothers and fathers are able to work full-time. Formal care at market prices is costly and affects the budget constraint so that it often absorbs a high share of the net gains of employment, reducing the utility of gainful employment. By contrast, subsidized public childcare reduces opportunity costs by lowering childcare costs. In the structural model of welfare state changes (Esping-Andersen 2002, 2009), subsidized public childcare is a crucial element in the shift from a familistic to a defamilialistic policy regime, which adapts to (most) women's occupational aspirations in OECD countries and allows women to better realize their employment preferences. Following the economic model of the family, this results in a higher utilization of public childcare and an increase in maternal employment (Blau 2003). The higher the subsidies and the more hours of formal care available, the higher the effect. The effects of subsidized childcare are relevant to high-income families with high earning power as well as low-income families, which could become economically self-sufficient (Blau 2003).

The link between subsidized childcare and increased female employment has been tested in various studies (for a detailed overview of the literature see, e.g., Morrissey 2017; Zoch and Hondralis 2017). Many empirical studies identify a positive effect of early childcare on parental labor force participation, but also demonstrate that findings vary by individual characteristics and context (Cascio, Haider, and Nielsen 2015; Morrissey 2017). At the individual level, the effects seem more pronounced for single mothers (Morrissey 2017), mothers with more children (Givord and Marbot 2014; Nollenberger and Rodriguez-Planas 2015; Zoch and Hondralis 2017), and those with lower incomes (Geyer, Haan, and Wrohlich 2015). At the contextual level, Zoch and Hrondalis (2017) find that childcare expansion has had a strong effect in West Germany but not in East Germany. Brilli, del Boca, and Pronzato (2016) find that the effect on maternal labor force participation is larger in regions with an initially lower number of available places. Moreover, benefits from childcare expansion are stronger in contexts with historically low female employment levels (Cascio, Haider, and Nielsen 2015; Nollenberger and Rodriguez-Planas 2015), or depend on the opportunities for informal childcare arrangements (e.g., Asai, Kambayashi, and Yamaguchi 2015). Overall, previous research shows a strong heterogeneity in the link between female labor force participation and public childcare arrangements.

Since 2006 the German Statistical Office has provided data on enrollment in subsidized childcare for toddlers aged 0–2 and for older children aged 3–5 (Hüsken 2011). These data are provided per year, but only as macro data aggregated at the county level. Indicators of overall enrollment and the share of full-time places have been available since 2007. Using this source, studies usually link county information on childcare coverage with microdata drawn from other surveys; e.g., the Micro Census

(MC) or the Socio-Economic Panel (SOEP). Using fixed effect models or other difference-in-difference approaches, researchers usually exploit the 'within' variance, often distinguishing between overall childcare places and full-time places. Using microdata from SOEP and event history models, Zoch and Hondralis (2017) find that higher childcare coverage (overall ratio) levels are modestly associated with shorter employment interruptions for mothers. Boll and Lagemann (2019) merge childcare data with the MC and use a multilevel panel model at the federal state level from 2006 to 2014 to show that increased childcare places for children aged 0–2 result in an increase in mothers' weekly working hours (intensive margin), while the extensive margin, FLP, is virtually unaffected. However, they only find significant effects for changes in the overall enrollment rate, and find that various policy indicators (e.g., the introduction of legal entitlement to a place in a day-care center from age 1) have rather weak effects when state-fixed effects are introduced into their specification. Müller and Wrohlich (2020), using childcare data and the MC and multi-level panel models with county fixed effects for the time span 2007 to 2014, find that a 1% increase in childcare coverage for children under age 3 increases the FLP (extensive margin) rates of mothers by 0.2 percentage points. They further find that this increase can mainly be attributed to an increase in part-time employment. Distinguishing between changes in full-time childcare places and overall coverage, the results indicate that the effect is driven by changes in the overall childcare coverage rates only, while the share of full-time places is insignificant. However, the overall effect becomes insignificant when controlling for county-specific time trends, while the effect of full-time care net of county-specific time trends is not reported (Müller and Wrohlich 2020, Table A-5). Zoch (2020), by combining the childcare data with the SOEP and drawing on multinomial models with an entropy matching procedure, reports a positive relationship between the availability of full-time places for toddlers and maternal employment.

All the studies mentioned above focus on childcare for toddlers aged 0–2, as the expansion of these places was the focus of recent policies. Contemporary research on the kindergarten age range 3–5 that utilizes childcare data from the German Statistical Office is rather sparse. Research on childcare expansion in Germany for the age group 3–5 tends to focus on the late 1990s, as the legal entitlement to a place in kindergarten was introduced in 1996. For instance, based on differences-in-differences analyses, Bauernschuster and Schlotter (2015) provide evidence that the introduction of this entitlement in 1996 had a positive effect on FLP.

In conclusion, the results suggest that increased childcare provision increases FLP, but it remains rather unclear whether overall enrollment or the share of full-time places is more important for increasing FLP, especially regarding childcare for toddlers aged 0–2. However, it seems reasonable that full-time childcare coverage allows for more flexible non-care-related time use and thus has a more beneficial effect on FLP.

Accordingly, research in Germany has identified particularly strong effects of full-time care on FLP in general (Boll and Lagemann 2019; Büchel and Spieß 2002). Similarly, full-time care was found to increase the likelihood of full-time employment (Zoch 2020) and part-time employment (20–35 hours) in West Germany (Müller and Wrohlich 2020). Even though the relative importance (in terms of effect magnitude) of full-time care compared to general care provision is not without question (e.g., Boll and Lagemann 2019; Müller and Wrohlich 2020), we assume that, in general, full-time places make it easier to reconcile work and family, and offer greater labor market opportunities than general care provision.

Given these results and regarding the general direction of effects, the first steps of the present analysis is to replicate existing research with new data and extend it further to childcare for children aged 3–5. Hence, we hypothesize:

H1: An increase in any kind of childcare coverage rate has a positive effect on FLP.

H2: An increase in the coverage rate of any kind of full-time childcare has an additional effect on FLP above increasing childcare rates per se.

Hypotheses H1 and H2 are well established and discussed and are partly supported by the literature above. One contribution of our paper is to replicate these findings, which were largely located at the micro level with macro data (Freese and Peterson 2017), and test to what extent these results are influenced by observation-specific time trends.

2.3 Effect heterogeneity in different area types and selection on trends

A major advantage in the previous literature has been the shift from a between-units comparison to a within-unit comparison. The latter ensures that unmeasured time-constant differences between the units of observation (e.g., counties) are controlled for and no longer bias the estimates. For instance, it is well known that due to historical path dependencies, childcare provision and FLP are higher in East Germany, which imposes a positive correlation between these two factors. In addition, German federal states with large Catholic populations tend to have fewer childcare places for children aged 0–2, while long-standing social democratic governments are associated with high childcare coverage and higher FLP (Andronescu and Carnes 2015). Several unobserved regional characteristics might thus induce a spurious association between childcare arrangements and FLP. To overcome this problem, many studies employ fixed-effects models, which account for all time-constant differences between observations by ‘de-meaning’ the data, hence controlling for any unobserved differences or unit-specific characteristics.

However, in many cases it has been recognized that unobserved differences or unit-specific characteristics can affect not only the time-constant levels but also time-varying trends (Dynarski, Jacob, and Kreisman 2018; Kneip and Bauer 2009; Ludwig and Brüderl 2018; Noelke 2016). For instance, Müller and Wrohlich (2020) state: "A crucial assumption for the causal interpretation of these estimates is that there are no region-specific time trends that drive both mothers' labor supply and the expansion of subsidized childcare." Supplementary analyses indicate that the effect of childcare coverage is not distinguishable from zero once region-specific time trends are controlled for in the FE models. It is not clear, however, whether this approach can account for all confounding factors due to trends, as this estimator also relies on the individual between-person variance within counties. A straightforward way of controlling for correlated trends in panel regression are fixed effects individual slope (FEIS) models (e.g., Brüderl and Ludwig 2015; Rüttenauer and Ludwig 2020; Wooldridge 2015; see method section). In the following, we discuss the issue of selection on trends in the case of childcare and FLP.

Our first argument for paying attention to these trends is straightforward: the existing literature already hints at biases due to selection issues. Müller and Wrohlich (2020) report a bias due to a positive selection of trends in West Germany, as they find an upward-biased estimate of the effect of childcare provision on FLP in FE models. Not only the historical divide between East and West but also the differences between German regions (Andronescu and Carnes 2015) can affect the relation between FLP and childcare. Daniel et al. (2019) point to important distinctions across different regional types such as rural and urban areas, a distinction that has barely been addressed in research to date. In the following we therefore derive possible theoretical explanations for the heterogeneous effects of childcare coverage according to the degree of urbanization. However, not only the effect of childcare on FLP but also time trends could vary across regional areas. Therefore, we also discuss potential reasons for regional heterogeneity.

It is known that childcare usage and demand are higher in Germany's more urban regions (Alt et al. 2017; Hubert, Lippert, and Alt 2019). While children in rural areas are more likely to live close to care-giving grandparents (García-Morán and Kuehn 2017), grandparental care is known to rarely allow for constant part-time or even full-time employment. Only 11% of German grandparents provide intense (24.7 hours a week on average) grandchild care (Di Gessa et al. 2016). Grandchild care is rather considered as supplementing professional childcare, and a huge majority of families in Germany has to rely on other forms of formal or informal childcare. While grandchild care can be assumed to play a larger role in more rural areas where the parties involved live nearer to each other, working mothers in urban areas have to rely more on formal childcare.

Moreover, research has shown that commuting distances have a pronounced negative effect on mothers' labor force supply, and thus, on average, mothers take jobs

with lower commute times than those of their male counterparts (Chidambaram and Scheiner 2020; Kawabata and Abe 2018). The negative association between commuting distance and labor supply is most pronounced in mothers with young children (Farré, Jofre-Monseny, and Torrecillas 2020; Kawabata and Abe 2018). Longer commuting times add to the time parents require their children to be in formal care (Daniel et al. 2019), thereby increasing the need for long-hours childcare. This indicates that childcare and housework at least partly constrain female labor force opportunities by reducing the spatial mobility and flexibility of mothers, and these constraints are most severe for mothers with young children. Following this strand of reasoning, we would expect an increase in the share of full-time places, especially for children aged 0–2, to have more positive effects on FLP in more rural areas, as required commuting times increase with the level of rurality (Dauth and Haller 2020). Based on these arguments, we derive the following hypothesis:

H3: Increases in general childcare provision have stronger effects on FLP in more urban regions. However, the share of full-time places has larger effects in more rural areas because this arrangement allows longer commuting times.

With regard to regional differences in time trends, our considerations are based on very simple assumptions. In Germany the childcare market is known to be a supply market: if childcare places exist, they are also used (Kreyenfeld and Hank 2000). Recent reports still show an unmet demand of 26.7% for childcare places for children aged 0–2 and of 3.3% for children aged 3–5. The supply gap is larger in West Germany, but no single federal state is managing to meet the need (BMFSFJ 2018). Hence, it is common in the literature to "use actual enrolment as a measure of available day care slots" (Jessen, Schmitz, and Waights 2020: 12) and to assume a "full take-up rate of newly established childcare slots" (Boll and Lagemann 2019: 3). Overall, the literature suggests that the actual childcare enrollment rates can be interpreted as a proxy for the supply of available childcare places. FLP, on the other hand, can be interpreted as a proxy for the demand for childcare. Hence, the highest demand for childcare can be expected in counties where FLP has a clear positive trend but childcare enrollment rates are comparatively low. A positive selection on trends, as reported by Müller and Wrohlich (2020), can occur if, for example, providers specifically select into these counties and expand their offerings there more frequently. Based on this assumption, we assume that estimates for areas where these conditions are present are more likely to be biased. Furthermore, as we know that unmet demand for childcare places is higher in the age group 0–2, we expect the bias due to a positive selection in trends in FE models to be higher in estimates of the effects of childcare coverage for children aged 0–2. Based on these arguments, we derive the following hypothesis:

H4: The effects of childcare coverage (especially for children aged 0–2) are more upwardly biased in conventional FE models in regions with increasing FLP and low starting levels of childcare.

In what follows, we first discuss the methods used in the analysis and then provide descriptive findings from our data and discuss results from FE and FEIS models. We further test for possible selection bias and show East–West differences in the effects of childcare expansion on FLP. By further differentiating between region types while also controlling for the possibility of correlated time trends, we seek to provide a detailed account of the link between increases in childcare and changes in female employment and to offer new insights into how the general association varies across different regional contexts.

3. Empirical approach

3.1 Data

To investigate the impact of childcare provision on FLP at the societal level, this paper uses macro data at the county level for all 401 German counties. We use data from Indicators and Maps on Spatial Development (INKAR, see Helmcke 2008), provided by the German Federal Office for Building and Regional Planning. The data comprise annual information from different data sources. The indicators used in this study are described below. To a great extent, information on childcare coverage has been available in Germany at the county level on an annual basis since 2007. Childcare data is available for different age groups. The indicators for childcare coverage are defined as the percentage of all children within the corresponding age group enrolled in subsidized childcare per county per year. We can utilize this information differentiated by age groups 0–2 and 3–5 (childcare coverage 0–2, childcare coverage 3–5). Further, we know the share of full-time places among all childcare places per age group (with a care time of 7 and more hours per day, share of full-time places 0–2, share of full-time places 3–5). Hence, we know the proportion of children enrolled per age group, and the proportion of full-time places. These indicators are currently only available until 2017.

To measure FLP, we use the percentage of women aged 20 to 45 who are employed in jobs that are subject to compulsory social insurance as provided by the employment statistics provided by the German Federal Employment Agency. That number includes trainees, part-time employees, interns, and working students. Also included are the marginally employed with low wages or short-term employment, as well as so-called mini-jobs. Not included are civil servants, the self-employed, family members helping

out, professional and temporary soldiers, and persons doing military or civilian service. This variable covers roughly 70% of the working population.

Note that we only have information on childcare coverage rates, rather than available places. As discussed above, we use the actual enrollment rate as a measure of available day-care places (Jessen, Schmitz, and Waights 2020). Nevertheless, we control for the demand side by adding the number of children aged 0–2, 4–6, and 6–18, as well as the fertility rate, as increasing numbers of children lower the rates of enrolled children when the number of childcare places remains unchanged. Further, we control for changes in the proportion of foreigners per county, to reflect events such as the 2015 migration crisis.

Moreover, our dependent variable measures total FLP. The age structure of the female population might thus influence the labor force participation rate, as well as the rate of childcare provision. Thus, we add the percentage of women aged 18–24, the percentage of women aged 25–64, and the percentage of women aged 65 and older as control variables. Further, we control for a wide range of educational and economic indicators that might be important in the current labor market situation and thus influence the rate of FLP. These include the percentage of people who have a university entrance qualification, the percentage of people leaving school without a certificate, the percentage of students, and the percentage of female students. Additionally, we control for the number of apprentice positions, GDP, the share of female employees in occupations subject to social security, and the number of female employees per 1000, the latter two to control for changes in the local labor market that affect women in particular.

All information is available for every county in East and West Germany for the period 2007 to 2017. We are only missing data on childcare for a very small number of counties in some years: Kleve, Mettman, and Wesel counties have missing data in 2017 and Mühldorf am Inn has missing data in 2015. We end up with 4,407 county years for 401 counties and have an almost full balanced macro panel for Germany, without attrition and with only 4 missing case-years. Further, we have 176 county years with coverage rates slightly over 100% for 3–5-year-olds, which are all set to 100%.[4]

To analyze regional effect heterogeneity, we utilize the well-known East–West difference as well as a county typology provided by INKAR to distinguish different area types, namely independent metropolitan counties (metro), urban counties (urban), agglomerated rural counties (agglomeration), and sparsely populated rural counties (rural).[5] This differentiation captures regional heterogeneity regarding population density

[4] This bias results from fuzzy areas in the data concerning age groups used to calculate the percentages, children using care arrangements in neighboring counties, and refugee children registered in the institution but not in the local data on the number of children.

[5] *Metropolitan counties*: Urban counties with at least 100,000 inhabitants. *Urban counties*: Counties with a population share in large and medium-sized cities of at least 50% and a population density of at least 150 inhabitants/km², as well as counties with a population density without large or medium-sized cities of at least

and labor market options and has been used by other researchers (e.g., Boll and Lagemann 2019). Descriptive information can be found in Figure A-1–A-4 and Table A-5, A-8–A-13 of the Online Appendix.

3.2 Fixed effect individual slope (FEIS) models

We apply FEIS models (Brüderl and Ludwig 2015; Rüttenauer and Ludwig 2020; Wooldridge 2015), which allow for unit-specific differences in levels (equal to conventional FE) and also for unit-specific slopes or trends over time. In conventional FE modelling,

$$\ddot{y}_i = \ddot{X}_i\beta + \ddot{\varepsilon}_i \qquad (1)$$

for each $i = 1, ..., N$ areas or units, where the $T \times 1$ vector $\ddot{y}_i = y_i - \bar{y}_i$, each column vector of $\ddot{X}_i$ is $\ddot{x}_{ik} = x_{ik} - \bar{x}_{ik}$ and $\ddot{\varepsilon}_i = \varepsilon_i - \bar{\varepsilon}_i$, with $\bar{\boldsymbol{y}}_i$, $\bar{\boldsymbol{x}}_{ik}$, $\bar{\boldsymbol{\varepsilon}}_i$ being the mean values of $\boldsymbol{y}_i$, $\boldsymbol{x}_{ik}$, and $\boldsymbol{\varepsilon}_i$ respectively.

FEIS models, by contrast, do not only 'de-mean' but also 'de-trend' the data by subtracting the predicted time trend for each individual observation. Formally, the model is given by:

$$\tilde{\boldsymbol{y}}_i = \tilde{\boldsymbol{X}}_i\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}_i \qquad (2)$$

for each i = 1, ..., N areas or units, where the $T \times 1$ vector $\tilde{\boldsymbol{y}}_i = \boldsymbol{y}_i - \hat{\boldsymbol{y}}_i$, each column vector of $\tilde{\boldsymbol{X}}_i$ is $\tilde{\boldsymbol{x}}_{ik} = \boldsymbol{x}_{ik} - \hat{\boldsymbol{x}}_{ik}$, and $\tilde{\boldsymbol{\varepsilon}}_i = \boldsymbol{\varepsilon}_i - \hat{\boldsymbol{\varepsilon}}_i$, with $\hat{\boldsymbol{y}}_i$, $\hat{\boldsymbol{x}}_{ik}$, $\hat{\boldsymbol{\varepsilon}}_i$ being the predicted values of $\boldsymbol{y}_i$, $\boldsymbol{x}_{ik}$, and $\boldsymbol{\varepsilon}_i$ respectively. The prediction is based on several pre-defined slope variables, such as time. Similar results can be obtained by including an interaction

150 inhabitants/km². *Rural counties with agglomeration approaches*: Counties with a population share in large and medium-sized cities of at least 50% but a population density below 150 inhabitants/km², as well as counties with a population share in large and medium-sized cities below 50% with a population density without large or medium-sized cities of at least 100 inhabitants/km². *Sparsely populated rural counties*: Counties with a population share in large and medium-sized cities below 50% and a population density excluding large and medium-sized cities below 100 inhabitants/km².
(Source: https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/kreise/siedlungsstrukturelle-kreistypen/kreistypen.html, accessed January 15, 2022)

between unit-specific dummies and the slope variables (e.g., Rüttenauer and Ludwig 2020; Wooldridge 2015). In our case we model the heterogeneous slopes by year and year squared. This seems to fit the time trends in female labor force participation (Figure A-1). Adding higher polynomials does not improve performance. We further add dummy variables for the economic recession in 2008/2009 and for the migration crisis in 2015, as the inclusion of time trends does not substitute completely for the year-fixed effects in the FE model.

As outlined earlier, conventional FE models rely on the assumption that time trends run parallel in 'treatment' and 'control' groups. Consequently, a correlation between the steepness or slope of the time trend of female labor force participation and the 'treatment' of new childcare provision would bias the estimates of conventional FE models. It is thus crucial to empirically test the possibility that time trends are correlated with the variable of interest. Rüttenauer and Ludwig (2020) propose an Artificial Regression Test, which constitutes an extension of the well-known Hausman test, and is able to detect confounding due to slope heterogeneity. By applying this test to the conventional FE models, we can investigate whether FE suffers from bias due to heterogeneous slopes. This matter is further discussed in the results section.

4. Results

4.1 Descriptive findings

In the following, we provide a brief discussion of the empirical distribution of our main variables, FLP and childcare coverage (see Figures A-1 to A-4 and Tables A-1, A-4 to A-10). Concerning FLP, rates went up from 53.6% to 65.1% in Germany in the period under study, an increase of a remarkable 11.5 percentage points. It increased in both parts of Germany: in East Germany from 60.2% to 72.5%, and in West Germany from 52.0% to 63.3%. Rural areas reported the steepest increase (12.9 percentage points), while the lowest increase was observed in metropolitan areas, with only 9.4 percentage points. On average, all groups have a more or less strong increase; however, there is broad heterogeneity within these groups at the county level.

The rates of childcare coverage 0–2 more than doubled from 15.7% in 2007 to 32.3% in 2017. The overall level was much higher in East Germany than West Germany (54.3% compared to 27.1% in 2017), and the increase was much steeper in West Germany (17.6 compared to 12.4 percentage points). The 2017 levels were highest in rural areas (36.2%) and lowest in urban areas (28.7%), the latter also reporting the steepest increase of 17.6 percentage points (see Figure A-2).

For the share of full-time places for children aged 0–2, we observe an even stronger increase from 7.3% to 17.4% in the period under study; hence rates more than doubled in 10 years. In West Germany rates more than quadrupled, from 2.5% in 2007 to 11.9% in 2017. In East Germany we observed an increase from 27.7% to 43.8%. Rates in the different regional area types were highest in metropolitan and rural areas (20.7% and 20.1% in 2017), and lower in agglomeration and rural areas (17.5% and 13.6% in 2017, see Figure A-2).

The overall picture for childcare for children aged 3–5 is much more homogeneous. East and West Germany report 90.6% and 93.4% in 2017. The change scores, when compared with 2007, are rather marginal and even negative. The greatest change was in urban areas, where enrollment rates for childcare for children aged 3–5 decreased by 1.2 percentage points.

Concerning the share of full-time places for children aged 3–5, however, we observe a steep increase overall and for all subgroups (see Figure A-4). While rates in West Germany more than doubled with an increase of 18.1 percentage points, from 16.3% in 2007 to 34.4% in 2017, East Germany started at a higher level (62% in 2007) and increased by only 14.4 percentage points to 76.4% in 2017. Regarding regional type, metropolitan areas have the highest levels (33.2% in 2007 and 53% in 2017) and the steepest increase (19.8 percentage points). The latter is almost as high in urban areas (19.7 percentage points), but urban areas report only 18.4% in 2007 and 38.1% in 2017. Hence, the largest increase in relative numbers was observed in urban areas. Readers may be surprised to learn that rural regions in fact do report the highest FLP and childcare enrollment rates, while urban regions tend to report the lowest. This is mainly due to the fact that many rural regions are located in East Germany, while many urban regions are located in West Germany (See Figure A-5). Here, the difference between East and West is reflected in these figures and it becomes clear how much Germany is still shaped by its history 30 years after reunification.

Overall, the differences in the level of and the increase in FLP seem rather marginal, with the exception of the known East–West divide in FLP. The steepest increases in childcare were in the share of full-time places 0–2 in West Germany, which increased by a factor of 4.5 (change rate, see Table A-1), followed by the share of full-time places 0–2 in urban areas (change rate 3.5). Also remarkable was the increase in childcare enrollment 0–2 in West Germany (change rate 2.9) and in urban areas (change rate 2.6). These indicators started from comparatively low levels in 2007.

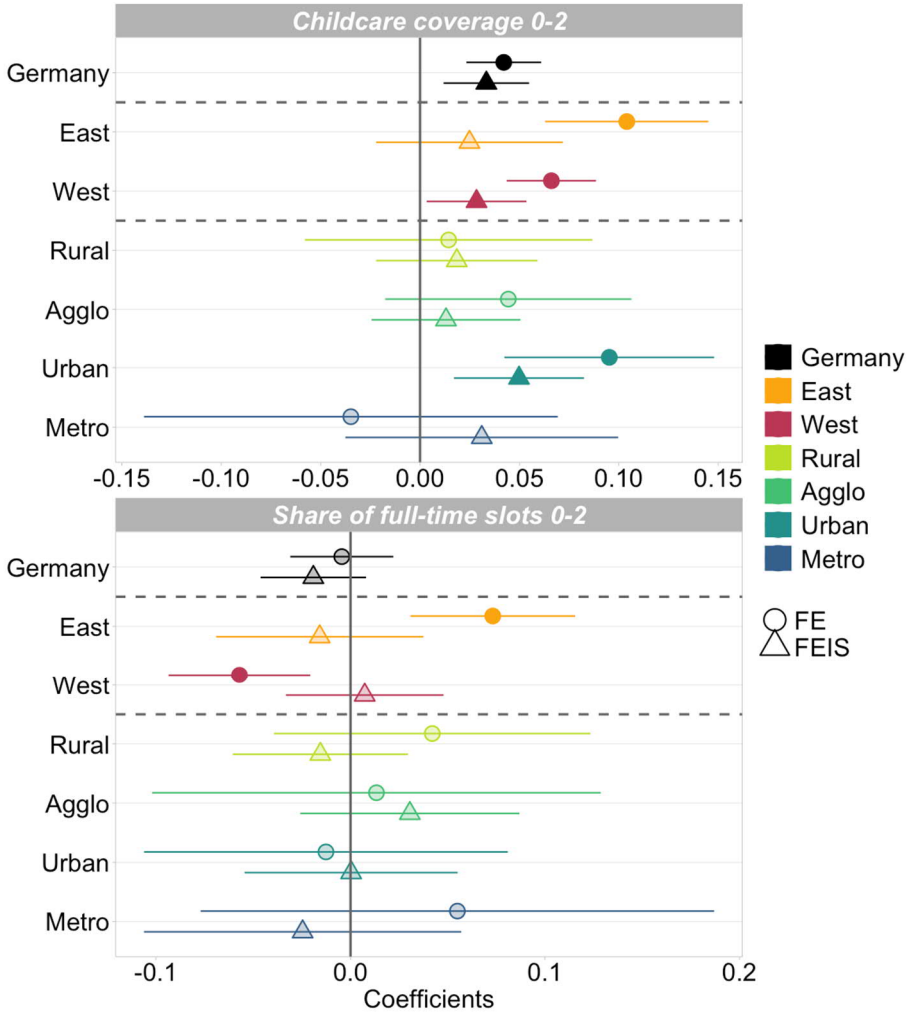
4.2 Fixed Effect (FE) and Fixed Effect Individual Slope (FEIS) models

Figure 1 contains forest plots of FE and FEIS estimates for all types of childcare coverage 0–2 and share of full-time places 0–2, overall, for East and West Germany, and for the 4 region types. Point estimates from FE models are indicated with dots, FEIS estimates with triangles. Error bars indicate 95% confidence intervals. All estimates with colored filling differ significantly from zero at the 5% level. Corresponding models for the estimates shown in the Figures can be found in the Online Appendix (Table A-2–A-3). The results are discussed in the order presented in the legend of Figure 1: hence, at the top, the black triangle and dot represent the effects of an increase in the childcare rate for children aged 0–2 on FLP from estimations with FE (triangle) and FEIS (dot) from a model including all regions of Germany (GERMANY). Note that an estimate of .1 means that a 1 percentage-point increase in childcare enrollment translates into a .1 percentage-point increase in FLP. The magnitudes of our effects are not directly comparable to the estimates provided by other researchers using individual data, as our dependent variable also includes women without children and hence presumably provides smaller effects for childcare enrollment.

First, considering the effects of changes in the rate of childcare coverage 0–2 on FLP in the model for Germany overall, we find that changes in coverage go hand in hand with rather small increases in FLP in the overall model in both FE (.04) and FEIS (.03) models, with the difference between FE and FEIS being rather marginal. Hence, when controlling for trends in FEIS, the effect is still statistically distinguishable from zero. In East Germany we find a comparatively large effect of .1 when using an FE framework, but the effect is considerably smaller (.02) and not distinguishable from zero when using an FEIS setup. By contrast, the effect remains statistically significant in West Germany when controlling for individual time tends, but the effect size is relatively small, with 0.03 percentage points.

When it comes to regional type, the only effects that are statistically distinguishable from zero are for urban regions, with the FE effect being much stronger than the FEIS (FE .1; FEIS .05). The effect estimates for rural, agglomeration, and metropolitan areas are rather small and not statistically distinguishable from zero in either FE or FEIS. The overall pattern of FEIS estimates points to a somewhat stronger effect in more urban areas. The Artificial Regression test (See Table 1) clearly hints that especially the models for East Germany, West Germany, and urban areas suffer from heterogeneous trends. As the effect of FEIS is smaller than that of FE, the selection bias stems from a positive selection in trends. Overall, we find reliable effects of increases in childcare coverage 0–2 on FLP in the models for Germany as a whole, West Germany, and urban areas, the strength of all of which tends to be overestimated if selection on trends is not controlled for. Especially in East Germany, the results of the FE models are largely influenced by selection on trends.

Figure 1: Forest plot of childcare enrollment effects for age group 0–2



Note: All models contain covariates and time and county fixed effects. Full models with all coefficients can be found below in Tables A-2 and A-3. All FEIS models contain quadratic terms for county time trends. Tests for selection on trends can be found in Table 1. Symbols for effects whose 95% confidence intervals (*p < 0.05) do not include the black vertical zero line are opaque, others are transparent; own calculations.
Source: German Statistical Office: Data on subsidized childcare; Federal Office for Building and Regional Planning (Indikatoren und Karten zur Raumentwicklung, INKAR): Indicators and Maps on the Spatial Development; German Federal Employment Agency: Data on female labor force participation.

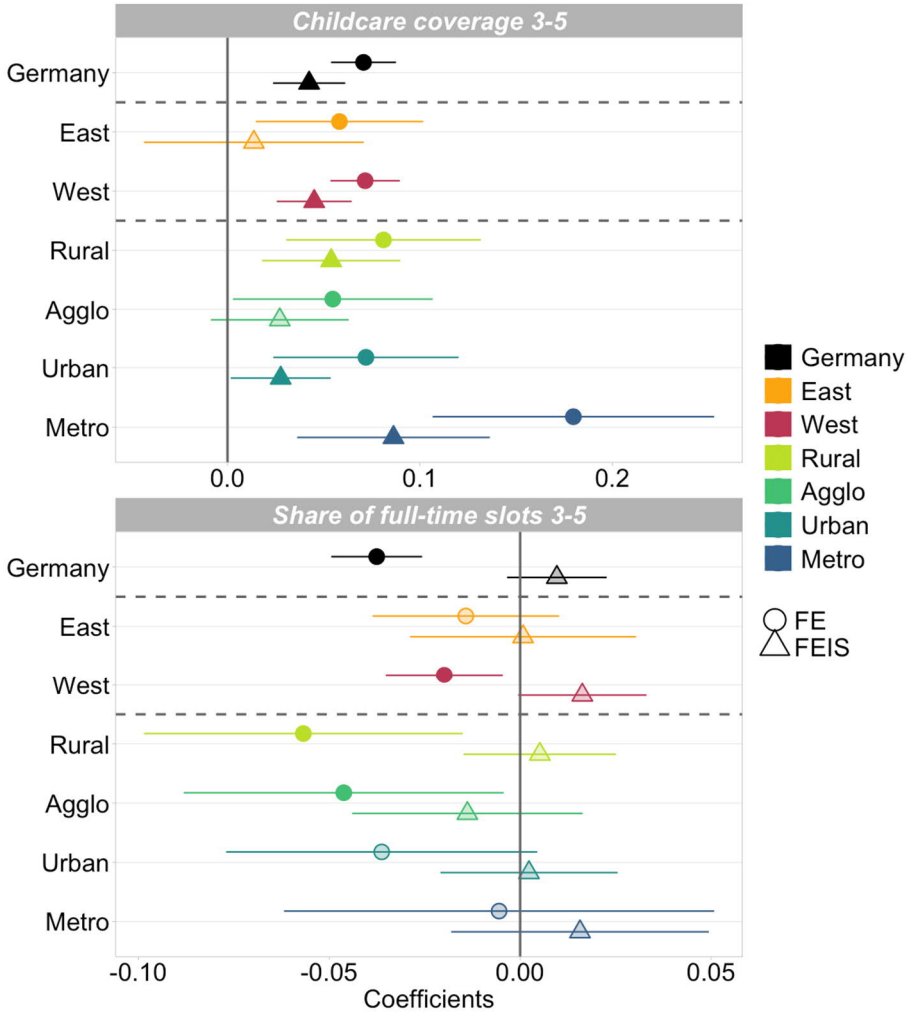
Table 1: Artificial regression tests for childcare for children aged 0–2

	χ^2	df	$P(>\chi^2)$	Area	Child care type
1	0.51	1.00	0.47	Germany	Childcare coverage 0 to 2
2	2.60	1.00	0.11	East	Childcare coverage 0 to 2
3	3.22	1.00	0.07	West	Childcare coverage 0 to 2
4	0.02	1.00	0.89	Rural	Childcare coverage 0 to 2
5	0.81	1.00	0.37	Agglo	Childcare coverage 0 to 2
6	3.98	1.00	0.05	Urban	Childcare coverage 0 to 2
7	0.91	1.00	0.34	Metro	Childcare coverage 0 to 2
8	0.11	1.00	0.75	Germany	Share of full-time places 0 to 2
9	3.73	1.00	0.05	East	Share of full-time places 0 to 2
10	2.98	1.00	0.08	West	Share of full-time places 0 to 2
11	1.19	1.00	0.27	Rural	Share of full-time places 0 to 2
12	0.02	1.00	0.90	Agglo	Share of full-time places 0 to 2
13	0.19	1.00	0.66	Urban	Share of full-time places 0 to 2
14	0.74	1.00	0.39	Metro	Share of full-time places 0 to 2

Source: Own calculations. See Figure 1 and Tables A-2 and A-3 for corresponding models.

The case is different for the share of full-time places for children aged 0–2. Here, the overall model for Germany and the area-type models provide only very small estimates in both FE and FEIS setups and neither of them is statistically distinguishable from zero. Only the FE estimates for East and West Germany are of considerable size, but both shrink toward zero when selection on trends is taken into account. Hence, when controlling for trends, none of the effect estimates for the share of full-time places for children aged 0–2 is of substantive magnitude. When taking the results from the Artificial Regression tests into account (see Table 1), we find confirmation that the estimates in the FE model for West and East Germany are likely to be biased.

Figure 2: Forest plot of childcare enrollment effects for age group 3–5



Note: All models contain covariates and time and county fixed effects. Full models with all coefficients can be found below in Tables A-2 and A-3. All FEIS models contain quadratic terms for county time trends. Tests for selection on trends can be found in Table 2. Symbols for effects whose 95% confidence intervals (*p < 0.05) do not include the black vertical zero line are opaque, others are transparent; own calculations.
Source: German Statistical Office: Data on subsidized childcare; Federal Office for Building and Regional Planning (Indikatoren und Karten zur Raumentwicklung, INKAR): Indicators and Maps on the Spatial Development; German Federal Employment Agency: Data on female labor force participation.

Table 2: Artificial regression tests for childcare for children aged 3–5

	χ^2	df	$P(>\chi^2)$	Area	Child care type
1	3.159026	1	0.075508	Germany	Childcare coverage 3 to 5
2	1.482081	1	0.223449	East	Childcare coverage 3 to 5
3	2.864475	1	0.090555	West	Childcare coverage 3 to 5
4	1.508792	1	0.219324	Rural	Childcare coverage 3 to 5
5	0.68294	1	0.408576	Agglo	Childcare coverage 3 to 5
6	2.494745	1	0.114227	Urban	Childcare coverage 3 to 5
7	9.825525	1	0.001721	Metro	Childcare coverage 3 to 5
8	12.98136	1	0.000315	Germany	Share of full-time places 3 to 5
9	0.986062	1	0.320707	East	Share of full-time places 3 to 5
10	3.895804	1	0.048407	West	Share of full-time places 3 to 5
11	7.581559	1	0.005897	Rural	Share of full-time places 3 to 5
12	2.162881	1	0.141379	Agglo	Share of full-time places 3 to 5
13	1.787341	1	0.181251	Urban	Share of full-time places 3 to 5
14	0.274139	1	0.600569	Metro	Share of full-time places 3 to 5

Source: own calculations. See Figure 2 and Tables A-2 and A-3 for corresponding models.

Turning to childcare coverage for children aged 3–5, we can observe a very clear pattern. All FE effects are positive and statistically distinguishable from zero and of considerable size (between .05 and .18) and all FEIS effects are smaller than FE effects, which shows that in all models the FE estimation suffers from a positive selection on trends. When controlling for this selection all estimates become noticeably smaller, but only the effect in East Germany and in agglomeration areas becomes statistically indistinguishable from zero. The strongest effects can be found in metropolitan areas (FEIS .09). Again, the difference between FEIS and FE estimates as well as the Artificial Regression tests (see Table 2) highlight the positive selection on trends.

For the proportion of the share of full-time places for children aged 3–5, the picture is quite the opposite. While all of the FE estimates except for metropolitan areas have a negative sign and some are of considerable magnitude – e.g., in the models for Germany, West Germany, and rural and agglomeration areas – the FEIS estimates all cluster around the zero line and none of them are statistically distinguishable from zero or reach considerable magnitudes. The slightly positive coefficients in FEIS and the negative coefficients in FE indicate that we have negative selection in this case: Those areas with a lower increase in FLP are more likely to expand their full-time childcare provision for

children 3–5. This is an interesting finding in itself. The Artificial Regression tests (see Table 2) further hint that this selection in more pronounced in West Germany and in rural areas.

5. Discussion

We investigate the effects of various childcare indicators on FLP in a macro panel data set containing data for all Germany in 2007–2017 using FE and FEIS models, taking into account possible biases due to selection on trends. Artificial regression tests show that we face issues with selection on trends, even when separating by East and West or by region type. Thus, we use both FE and FEIS models as a best practice approach when using macro data (e.g., Dynarski, Jacob, and Kreismann 2018). In this way we can assess the extent to which previous results have been determined by selection on trends and provide findings adjusted for these biases.

Regarding the effect of childcare 0–2 on FLP, our macro analysis supports previous findings: an increase in childcare provision correlates with an increase in FLP in West Germany (Müller and Wrohlich 2020), and this positive effect is to a certain extent determined by selection on different time trends. Our findings that FE provides considerably larger estimates (.07) and that these effects are considerably smaller when controlled for selection (.03, see Tables A-7, A-6) are in line with the supplementary analyses conducted by Müller and Wrohlich (2020, Table A-5). In their study, the effect of overall childcare coverage 0–2 in West Germany is reduced by more than 50% (from .22* to .09) after controlling for region-specific trends. This again highlights the presence of selection problems, and demonstrates that macro modelling approaches also have to take heterogeneous trends into account. The significance of selection problems is further evident in childcare enrollment in East Germany, as FEIS shows considerably smaller effect estimates than FE in East Germany. Ignoring the selection issue might lead researchers to report that the effect is strongest in East Germany, which is obviously not the case when selection bias is controlled for. In addition, and not reported in the existing literature so far, our spatially differentiated analysis reveals that increases in full-time care 0–2 mostly boosts FLP in urban areas – or to be precise, in urban areas in West Germany.

Regarding the share of full-time places 0–2, our results show that no effects are statistically distinguishable from zero in any model when selection on trends is taken into account. This conclusion contradicts results by Zoch (2020), who concludes that particularly increases in the share of full-time places are related to increases in female employment. However, the finding is in line with Müller and Wrohlich 2020, who also report an insignificant and negligible effect of the share of full-time places on FLP.

Anyway, this point must be interpreted with great caution, as our dependent variable for FLP also does not include the intensive margin of FLP and hence the magnitude of employment, but participation only. Here, a more detailed replication is necessary, which we cannot provide based on our macro data.

In the case of childcare coverage for children aged 3–5, our research provides new and previously unknown insights, as research on the relation between childcare coverage for the age group 3–5 and FLP is comparatively rare. We find a very robust and substantial positive effect of increases in childcare coverage for children aged 3–5 in almost all models, even if controlling for selection on trends considerably reduces the effect size. We find increases to be especially effective in West Germany and in rural and metropolitan areas.

Again, we do not find any relevant effects of changes in the rate of the share of full-time places for children aged 3–5 after controlling for selection on trends. This variable in particular shows how important it is to control for trends. If researchers fail to control for trends when estimating the effect of an increase in fulltime places for this age group, they will find negative effects which are statistically distinguishable from zero. The same holds true for the negative estimate of increases in the rate of the share of full-time places for children aged 0–2 on FLP in the FE model for West Germany. Not controlling for selection issues might produce significant but completely counterintuitive results: we cannot think of any theoretical reasons why an increase in the share of children using full-time childcare should lead to decreases in FLP.

Returning to our hypotheses, we can state that our analysis supports hypothesis H1. Increases in childcare coverage for children of all age groups go hand in hand with increases in the local rate of FLP, even if not everywhere. Regarding our Hypothesis H2, we do not find any evidence that increases in the share of full-time places have any additional effect when the overall rate is controlled for; hence H2 is not supported.

Regarding H3, we find evidence of notable effects for childcare coverage 0–2 in urban areas where the demand is higher. As our descriptive analyses shows, urban areas had the lowest share and the steepest increase in childcare places for children aged 0–2, which directly translated into increasing FLP. The lack of effect in more rural areas could be explained by the higher share of grandparental care, as increasing childcare rates in more rural areas could lead to a crowding out of alternative childcare arrangements without affecting FLP. Moreover, the positive effect of childcare for children aged 0–2 in urban areas might be connected to highly educated mothers living in suburban areas rather than in the metropolitan centers, as it is known that childcare for this age group is mainly used by higher-educated mothers (Schober and Spieß 2013). Further, urban areas and West Germany reported lower starting levels of childcare coverage for children aged 0–2. Hence, our findings are also in line with Brilli, del Boca, and Pronzato (2016), who

find that effects of childcare usage on maternal labor force participation are larger in regions with an initially lower number of available places.

Notable effects of enrollment of children aged 3–5 on FLP are found in West Germany and in rural areas, but are most pronounced in metropolitan areas, where we find by far the greatest effect in FEIS. The latter finding might be partly explained by the increasing scarcity of childcare places due to population growth in more urban areas, which has further increased due to recent immigration. However, to explain our findings further research is needed that takes an in-depth look at different population groups in different areas, which is beyond the scope of our paper.

Considering H4, we find the increase in FLP was relatively constant in all areas considered (see Table A-1). Strong increases in childcare for children aged 0–2 were observed in West Germany and in urban areas. Here, the artificial regression tests for the corresponding estimates (See Table 1) show a strikingly high probability of a (positive) selection on trends. Hence, considering childcare for children aged 0–2, our findings provide some evidence that FE estimates for areas with steeper increases are likely to be prone to selection issues, which supports Hypothesis 4. When it comes to the share of full-time places for children aged 0–2, the hypothesis does not hold true. Here, even steeper childcare increases were observed in urban areas in West Germany, but the former estimate suffers from a strong but negative selection on trends (as the FE estimate is smaller than the FEIS), and the latter estimate for urban areas is not biased at all. Regarding childcare coverage for children aged 3–5, we observe almost no changes in the rates (Table A-1), but we do find strong bias due to positive selection, especially in metropolitan areas (Table 1 and Figure 1). Considering full-time places for the age group 3–5, we observe increasing rates but negative selection in all estimates, especially those in the effects for Germany, West Germany, and rural areas.

Hence, our attempt at an explanation does not cover all possibilities and especially fails when it comes to negative selection, which could occur in at least two scenarios. First, day-care providers may invest in structurally deteriorating regions – possibly due to local political decisions. Second, an increasing FLP could go in hand with an increasing number of young children, which might lead to a decreasing overall rate of children in day care. Here, negative selection could suggest that the providers cannot meet the rapidly growing demand, although this should actually be taken into account by our models. We do not claim to thoroughly discuss the topic, but we are not aware of any research that deals with selection in day-care places and FLP and we would like to encourage further research in the area.

Our findings have political implications for childcare policy in Germany. First, our results suggest that increased enrollment rates 0–2 are especially effective in urban areas in West Germany, whereas rural or agglomeration areas with similar rates of FLP might react less to changes in childcare enrollment rates. Furthermore, we did not find any

substantial effect of increases in the share of full-time places, but this is to be interpreted with great caution. The lack of any results here is likely due to the lack of explanatory power regarding the extensive margin of our FLP variable. Second, the magnitude and stability of the effects of an increase in the enrollment rate of children aged 3–5 is striking. This age group plays a rather subordinate role in the current discussion, but our results clearly show that this subordinate role is not appropriate. Policymakers aiming to increase FLP should also pay particular attention to increasing the availability of kindergarten places.

Our study has several limitations. First, we are fully aware that our results are based on observational data. Though we extend previous findings by controlling for an additional source of bias, we cannot be sure that our data does not suffer from additional sources of endogeneity. Second, contrary to other studies that use rich individual data, our macro analysis lacks depth of field in terms of individual circumstances: We do not have information about individual women; e.g., the number of children they have or their educational background. Furthermore, we only have limited information on mothers' FLP. We are fully aware that detailed information about individual care needs and arrangements as well as the individual employment intensity and elasticity of mothers with young children would be desirable. Further research should thus aim to connect our finding of regionally heterogeneous effects and trends to individual data to gain a more comprehensive understanding of the micro mechanisms. For instance, it is necessary to test how far our results can be explained by heterogeneous effects of childcare on the intensive margins; e.g., part-time employment. While we provide some empirical evidence that spatial-effect heterogeneity exists in the case of childcare and female employment, we can only speculate on the detailed mechanisms underlying these regional differences. More detailed analyses are thus necessary before making far-reaching policy recommendations regarding these findings. We also checked if the data might suffer from spatial autocorrelation. As expected, the level of cross-sectional spatial autocorrelation is very high (e.g., average Moran's I of 0.61 for FLP and 0.86 for childcare below age 3). However, the de-trended data – as used in the FEIS models – exhibits much lower spatial autocorrelation (average Moran's I of 0.13 and 0.12 respectively). It is thus unlikely that our model estimates suffer from a noticeable bias due to autocorrelation (Rüttenauer 2019). Further work should nevertheless investigate spatial relations like spatial spillover in more detail.

Overall, the findings presented here underline the fact that the expansion of childcare caused an increase in FLP, and they enrich previous knowledge by differentiating childcare indicators, time trends, and regions, providing a possible explanation for the rather mixed results in the literature. We show that estimates differ greatly depending on different model specifications and regional differences. In particular, our results highlight the impact of childcare for toddlers in urban areas. Further, they highlight the important

but overlooked role of childcare for children aged 3–5. Because of the considerable differences between the FE and the FEIS results, we recommend that future research on the topic take heterogeneous levels and trends into account. Contrary to FE, our FEIS estimates did not provide large negative effects for any kind of childcare – a counter-intuitive result when relying on FE only. In our paper we show that an in-depth analysis with more advanced methods (FEIS) helps to shed light on seemingly counterintuitive results and the need for spatial analyses. Furthermore, our account of the regional heterogeneity highlights the varying effects of childcare on FLP across different regional types, and we hope our findings stimulate further research in this direction.

Bibliography

Alt, C., Gesell, D., Hubert, S., Hüsken, K., Kuhnke, R., and Lippert, K. (2017). DJI-Kinderbetreuungsreport 2017 [Childcare Report 2017]. München: Deutsches Jugendinstitut.

Andronescu, C.G. and Carnes, M.E. (2015). Value coalitions and policy change: The impact of gendered patterns of work, religion and partisanship on childcare policy across German states. *Journal of European Social Policy* 25(2): 159–174. doi:10.1177/0958928715573480.

Asai, Y., Kambayashi, R., and Yamaguchi, S. (2015). Childcare availability, household structure, and maternal employment. *Journal of the Japanese and International Economies* 38: 172–192. doi:10.1016/j.jjie.2015.05.009.

Bauernschuster, S. and Schlotter, M. (2015). Public child care and mother's labor supply: Evidence from two quasi-experiments. *Journal of Public Economics* 123: 1–16. doi:10.1016/j.jpubeco.2014.12.013.

Becker, G.S. (1991). *A treatise on the family*. Enlarged Edition. Cambridge, Mass.: Harvard University Press. doi:10.4159/9780674020665.

Blau, D.M. (2003). *The child care problem*. New York: Russel Sage Foundation.

BMFSFJ (2018). Kindertagesbetreuung ompakt: Ausbaustand und Bedarf 2017.

Boll, C. and Lagemann, A. (2019). Public childcare and maternal employment – New evidence for Germany. *Labour*. doi:10.1111/labr.12143.

Brilli, Y., Del Boca, D., and Pronzato, C.D. (2016). Does child care availability play a role in maternal employment and children's development? Evidence from Italy. *Review of Economics of the Household* 14(1): 27–51. doi:10.1007/s11150-013-9227-4.

Brüderl, J. and Ludwig, V. (2015). Fixed-effects panel regression. In: Best, H. and Wolf, C. (eds.). *The Sage handbook of regression analysis and causal inference*. Los Angeles.: Sage Reference: 327–357. doi:10.4135/9781446288146.n15.

Büchel, F. and Spieß, C.K. (2002). Kindertageseinrichtungen und Müttererwerbstätigkeit: Neue Ergebnisse zu einem bekannten Zusammenhang. *Vierteljahrshefte Zur Wirtschaftsforschung* 71(1): 95–113. doi:10.3790/vjh.71.1.95.

Cascio, E.U., Haider, S.J., and Nielsen, H.S. (2015). The effectiveness of policies that promote labor force participation of women with children: A collection of national studies. *Labour Economics* 36: 64–71. doi:10.1016/j.labeco.2015.08.002.

Chidambaram, B. and Scheiner, J. (2020). Understanding relative commuting within dual-earner couples in Germany. *Transportation Research Part A: Policy and Practice* 134(February): 113–129. doi:10.1016/j.tra.2020.02.006.

Daniel, H-D., Hannover, B., Köller, O., Lenzen, D., McElvany, N., Roßbach, H-G., Seidel, T., Tippelt, R., and Wößmann, L. (eds.) (2019). *Region und Bildung. Mythos Stadt - Land*. Münster: Waxmann.

Dauth, W. and Haller, P. (2020). Is there loss aversion in the trade-off between wages and commuting distances? *Regional Science and Urban Economics* 83 doi:10.1016/j.regsciurbeco.2020.103527.

Di Gessa, G., Glaser, K., Price, D., Ribe, E., and Tinker, A. (2016). What drives national differences in intensive grandparental childcare in Europe? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 71(1): 141–153. doi:10.1093/geronb/gbv007.

Dynarski, S., Jacob, B., and Kreisman, D. (2018). How important are fixed effects and time trends in estimating returns to schooling? *Journal of Applied Econometrics* 33(7): 1098–1108. doi:10.1002/jae.2653.

England, P. (2010). The gender revolution. *Gender and Society* 24(2): 149–166. doi:10.1177/0891243210361475.

Esping-Andersen, G. (2002). A child-centred social investment strategy. In: Esping-Andersen, G., Gallie, D., Hemerijk, A., and Myers, J. (eds.). *Why we need a new welfare state*. Oxford: Oxford University Press: 26–67. doi:10.1093/0199256438.003.0002.

Esping-Andersen, G. (2009). *Incomplete revolution: Adapting welfare states to women's new roles*. Cambridge: Polity Press.

Farré, L., Jofre-Monseny, J., and Torrecillas, J. (2020). Commuting time and the gender gap in labor market participation. (IZA Discussion Paper 13213). doi:10.2139/ssrn.3596679.

Freese, J. and Peterson, D. (2017). Replication in social science. *Annual Review of Sociology* 43: 147–165. doi:10.1146/annurev-soc-060116-053450.

Gangl, M. and Ziefle, A. (2015). The making of a good woman: Extended parental leave entitlements and mothers' work commitment in Germany. *American Journal of Sociology* 121(2): 511–563. doi:10.1086/682419.

García-Morán, E. and Kuehn, Z. (2017). With strings attached: Grandparent-provided child care and female labor market outcomes. *Review of Economic Dynamics* 23: 80–98. doi:10.1016/j.red.2016.09.004.

Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics* 36(C): 84–98. doi:10.1016/j.labeco.2015.07.001.

Givord, P. and Marbot, C. (2014). Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies. *Labour Economics* 36: 99–111. doi:10.1016/j.labeco.2015.07.003.

Helmcke, T. (2008). Kita vor Ort – Betreuungsatlas auf Ebene der Jugendamtsbezirke 2010. *Wirtschaft Und Statistik* 3: 207–216.

Hubert, S., Lippert, K., and Alt, C. (2019). Frühe Bildung Bilanz und Perspektiven für Deutschland: Unerfüllte Betreuungswünsche. *DJI Impulse* 121(1): 38–40.

Hüsken, K. (2011). Kita vor Ort – Betreuungsatlas auf Ebene der Jugendamtsbezirke 2010. München: Deutsches Jugendinstitut.

Immervoll, H. and Barber, D. (2006). Can parents afford to work? Childcare costs, tax-benefit policies and work incentives. (IZA Discussion Paper Series 1932). doi:10.2139/ssrn.878665.

Jessen, J., Schmitz, S., and Waights, S. (2020). Understanding day care enrollment gaps. *Journal of Public Economics* 190: 104252. doi:10.1016/j.jpubeco.2020.104252.

Kawabata, M. and Abe, Y. (2018). Intra-metropolitan spatial patterns of female labor force participation and commute times in Tokyo. *Regional Science and Urban Economics* 68: 291–303. doi:10.1016/j.regsciurbeco.2017.11.003.

Kneip, T. and Bauer, G. (2009). Did unilateral divorce laws raise divorce rates in Western Europe? *Journal of Marriage and Family* 71(3): 592–607. doi:10.1111/j.1741-3737.2009.00621.x.

Kreyenfeld, M. and Hank, K. (2000). Does the availability of childcare influence the employment of mothers? Findings from Western Germany. *Population Research and Policy Review* 19(4): 317–337. doi:10.1023/A:1026556309080.

Ludwig, V. and Brüderl, J. (2018). Is there a male marital wage premium? New evidence from the United States. *American Sociological Review* 83(4): 744–770. doi:10.1177/0003122418784909.

Morrissey, T.W. (2017). Child care and parent labor force participation: A review of the research literature. *Review of Economics of the Household* 15(1): 1–24. doi:10.1007/s11150-016-9331-3.

Müller, K.U. and Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare. *Labour Economics* 62: 101776. doi:10.1016/j.labeco.2019.101776.

Noelke, C. (2016). Employment protection legislation and the youth labour market. *European Sociological Review* 32(4): 471–485. doi:10.1093/esr/jcv088.

Nollenberger, N. and Rodriguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics* 36: 124–136. doi:10.1016/j.labeco.2015.02.008.

OECD (2017). OECD Employment Outlook.

Preisner, K., Neuberger, F., Bertogg, A., and Schaub, J.M. (2020). Closing the happiness gap: The decline of gendered parenthood norms and the increase in parental life satisfaction. *Gender and Society* 34(1): 31–55. doi:10.1177/0891243219869365.

Preisner, K., Neuberger, F., Posselt, L., and Kratz, F. (2018). Motherhood, employment, and life satisfaction: Trends in Germany between 1984 and 2015. *Journal of Marriage and Family* 80(5): 1107–1124. doi:10.1111/jomf.12518.

Rüttenauer, T. and Ludwig, V. (2020). Fixed effects individual slopes: Accounting and testing for heterogeneous effects in panel data or other multilevel models. *Sociological Methods and Research*, doi:10.1177/0049124120926211.

Rüttenauer, T. (2019). Spatial regression models: A systematic comparison of different model specifications using monte carlo experiments. *Sociological Methods and Research*. doi:10.1177/0049124119882467.

Schmitz, S., Spieß, C.K., and Stahl, J.F. (2017). Kindertageseinrichtungen: Ausgaben der Familien sind von 1996 bis 2015 mitunter deutlich gestiegen. *DIW Wochenbericht* (41): 889–903.

Schober, P.S. and Spieß, C.K. (2013). Early childhood education activities and care arrangements of disadvantaged children in Germany. *Child Indicators Research* 6(4): 709–735. doi:10.1007/s12187-013-9191-9.

Thévenon, O. (2013). Drivers of female labour force participation in the OECD. (OECD Social, Employment and Migration Working Papers 145). OECD Publishing.

Wooldridge, J.M. (2015). *Introductory econometrics: A modern approach*. Cengage Learning.

Zoch, G. (2020). Public childcare provision and employment participation of East and West German mothers with different educational backgrounds. *Journal of European Social Policy* 30(3): 370–385. doi:10.1177/0958928719892843.

Zoch, G. and Hondralis, I. (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(5): 693–707. doi:10.1093/esr/jcx068.

Zoch, G. and Schober, P.S. (2018). Public child-care expansion and changing gender ideologies of parents in Germany. *Journal of Marriage and Family* 80(4): 1020–1039. doi:10.1111/jomf.12486.