*Economics of Transition and Institutional Change*

WILEY

ECONOMICS OF TRANSITION AND INSTITUTIONAL CHANGE

ORIGINAL ARTICLE OPEN ACCESS

# The Heterogeneous Impact of Children on Maternal Employment: Evidence From East and West Germany

Johannes Köckeis[1,2] | Sven Stöwhase[1]

[1]Fraunhofer Institute for Applied Information Technology FIT, Sankt Augustin, Germany | [2]Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany

**Correspondence:** Johannes Köckeis (johannes.koeckeis@fit.fraunhofer.de)

**Received:** 26 November 2024 | **Revised:** 6 August 2025 | **Accepted:** 6 October 2025

**Keywords:** female employment | fertility | Germany

## ABSTRACT

This study investigates the causal effect of fertility on female labour market outcomes in East and West Germany. We use twin births as an exogenous variation for family size. Our results suggest a negative relationship between the number of children and maternal labour market outcomes. However, this connection is significantly stronger in West Germany than in East Germany for the second and third child. By the fourth child, the effects in the two parts of the country become more similar. Further subgroup analyses suggest that these differences can be explained by regional institutional conditions, such as the availability of public childcare facilities, and to a lesser extent by attitudes towards working mothers.

**JEL Classification:** D1, J13, J22, Z1

## 1 | Introduction

Although the labour force participation of women in OECD countries has increased significantly in recent decades, working-age men remain significantly more likely to be employed than women (Fluchtmann and Patrini 2023). Despite similar levels of education, gender gaps in wages and hours worked continue to exist (Olivetti and Petrongolo 2016). Previous literature has identified children as the main reason for these differences (Kleven et al. 2019).

A negative correlation between motherhood and labour force participation represents a challenge for ageing societies. As a significant decline in the working population is expected in many countries in the coming decades, fertility must remain high. However, the potential to attract more mothers to the labour market is very high (Hellwagner et al. 2022). This poses a dilemma for policymakers because policies to increase fertility should not, ceteris paribus, negatively affect maternal employment rates. Therefore, a better understanding of the connection between fertility and maternal labour market outcomes is important.

A major obstacle in the empirical analysis of this relationship is reverse causality. On the one hand, children can negatively affect female labour force participation by introducing constraints on labour market activity. On the other hand, mothers may self-select into larger families if they are less inclined to work or if their expected wage is low. The problem of identifying causal effects is typically solved by natural experiments that increase fertility. Several studies use the sex composition or sex preference of the first two children as an instrument for family size (J. Angrist and Evans 1998). Other studies use miscarriage or infertility as exogenous variation (Agüero and Marks 2011; Ashcraft et al. 2013). Cultural components, such as the fifth year of the Chinese lunar calendar (Vere 2008), or political interventions, such as the two-child policy in China (Wu 2022), are additional natural experiments that increase the number of children. In Lundborg et al. (2017), IVF treatment success is used as an exogenous variation for fertility. We apply a classic approach by using twin births as an instrument for exogenous variation in the number of children (Rosenzweig and Wolpin 1980; Bronars and Grogger 1994; J. Angrist and Evans 1998).

The effect of fertility on female labour force participation is usually large and negative in developed countries (Aaronson et al. 2021). Studies from several countries, particularly those of Western countries, have confirmed this relationship.[1] However, differences have also been identified. In a cross-country study, Baranowska-Rataj and Matysiak (2016) demonstrate that the magnitude of this relationship varies significantly among groups of European countries. This raises the question of the drivers of these heterogeneous effects.

A substantial proportion of the literature attributes this to the different institutional frameworks in these countries. Specifically, this encompasses family policy measures, such as parental leave and the provision of state-subsidised childcare, which facilitate the reconciliation of work and family life for mothers (Boeckmann et al. 2015; Kremer 2007).

Another potential factor is the influence of social norms on women's labour market activities. These norms are related to the societal expectations concerning the roles and responsibilities of men and women following the entry into parenthood. This literature argues that traditional gender role attitudes contribute to a more specialised division of labour among parents, which reduces mothers' employment and income (Pfau-Effinger 2012). These gender norms vary significantly across different countries (Kleven, Landais, Posch, et al. 2019).

The study of heterogeneous effects is especially interesting when examining regions that have been shaped by different political regimes. For instance, it has been demonstrated that state socialist regimes have influenced both gender role attitudes (Campa and Serafinelli 2019) and preferences for social policy (Alesina and Fuchs-Schündeln 2007). Due to the complexity of welfare states, the variations in social norms across regions within countries, and differences in economic circumstances, comparisons between Western countries and former socialist countries might be misleading.

Therefore, this study uses the unique historical setting of Germany, which was divided into a socialist and Western country for over 40 years. The family policies of the two countries showed enormous differences, particularly regarding women's employment. The German Democratic Republic (GDR) implemented a series of policies with the objective of promoting the employment of mothers. These policies included the expansion of public childcare and the introduction of paid parental leave (Trappe 1996). In contrast, the family policy of the Federal Republic of Germany (FRP) was strongly driven by the male breadwinner model, making it difficult for women to reconcile work and family life (Mätzke 2010).

After the reunification of Germany in 1990, a complete integration of the institutions of the GDR into the FRG was implemented. This also reflected the family policy. One exception was the provision of state-subsidised childcare for children under the age of three, which remained significantly higher in East than in West Germany. Furthermore, even after reunification, it can be observed that egalitarian gender roles remain far more prevalent in East Germany (Pfau-Effinger and Geissler 2002).

Our paper contributes to the literature on the influence of children on maternal labour market outcomes in East and West Germany (Jessen 2022; Collischon et al. 2020). Although previous studies identify cultural differences between East and West Germany as explanations for varying levels of child penalties, they contribute little to answering the question of how additional children affect the magnitude of the penalty and whether heterogeneity also exists in this dimension across the two regions. We contribute to this strand of literature by adding the causal effect of the number of children on maternal labour market outcomes. We show that differences between East and West Germany are not only driven by the transition into motherhood but also by heterogeneous effects depending on the number of children. Such differentiation allows for a more accurate understanding of how cultural and institutional contexts influence the relationship between fertility and female employment—an insight that is critical for designing targeted policy interventions aimed at promoting gender equality in the labour market. If cultural and institutional conditions differ across the two regions, policymakers may need to apply distinct instruments tailored to each context.

Our analysis is based on a unique administrative dataset, covering the whole universe of income taxpayers in Germany. We find a negative relationship between family size and maternal labour market outcomes in West and East Germany. Our OLS estimates overstate this negative effect, which indicates that women with lower inclination to work self-select into larger families. However, the relationship between fertility and maternal employment follows completely different patterns in East and West Germany. The effect of having a second or third child on a mother's employment is much stronger in West Germany (9.5% and 10.2% respectively) than in East Germany (1.6% and 3.0%, respectively). With the fourth child, the two parts of the country become more similar. A similar pattern emerges regarding the impact on maternal labour income. Large differences are also evident in the subanalysis according to the degree of urbanisation. While peripheral rural regions in West Germany show the strongest effects (between 9.4 and 20.3%), no statistically significant results can be determined for East Germany. Another contribution of our paper is the investigation of the mechanisms of this causal relationship by using information on the region around the place of residence. The subgroup analysis conducted for mothers with toddlers suggests that regional institutional conditions, such as the public provision of childcare, may account for a significant portion of the observed disparities between East and West Germany. The prevailing social norms regarding mothers' employment in the immediate area have a smaller, but still substantial effect.

The rest of the paper is organised as follows: the next section provides background information on fertility and labour supply of women in Germany, as well as an overview of the related literature. Section 3 describes the data used and presents descriptive statistics. Section 4 outlines our estimation strategy. Our results are presented in Section 5. Section 6 summarises and discusses our findings.

## 2 | Background

### 2.1 | Fertility, Labour Supply and Local Childcare Supply in East and West Germany

Even several decades after German reunification, substantial disparities persist between East and West Germany in numerous economic and social areas.[2] The differences in women's employment are particularly striking. Overall, the women's employment in Germany has increased significantly over the last 20 years. In 2002, it was only around 45%, but by 2018 it had gradually increased to around 56%. However, as the left-hand side of Figure 1 reveals, a persistently higher employment rate can be observed in East Germany over the entire period. While at the beginning of the observation period it was only around 3 pp (percentage points) higher than in West Germany, by 2018 this difference had increased to 5 pp. One explanation for the observable differences could be the different fertility rates in East and West Germany. The systemic upheaval in East Germany associated with the reunification of Germany led to considerable economic changes and uncertainties, which in turn led to a significant decline in fertility rates in East Germany in the following years (Witte and Wagner 1995). As can be seen from the right side of Figure 1, the total fertility rate in East Germany in 2002 was significantly lower at about 1.2 per woman than in West Germany just under 1.4 per woman. By 2008, the birth rate in East Germany had equalised with that in West Germany and has since increased at a steady rate in both parts. However, this convergence cannot be observed in the employment rates.

The link between fertility and female employment rates alone may be misleading, though, as the compatibility of family and work is largely determined by the availability and costs of childcare options. Traditionally, there have been significant differences in the availability of (publicly provided) childcare between East and West Germany. This particularly concerns the availability of childcare places for children up to the age of three. In 2018, the coverage rate for children under three years of age in East Germany was 51.5%. It was similarly high in previous years. By contrast, childcare options for children under three years old were poorly developed in West Germany until 2005. Following a legal change in 2008 (Kinderförderungsgesetz —KiFöG[3]) that established a legal entitlement to a childcare place for all children aged one to three, the coverage rates have increased in West Germany. In 2018, the coverage rate in West Germany was 29.4%, which is still around 22 pp lower than that in East Germany. These differences between East and West Germany are even more pronounced in full-time care (see Supporting Information S1: Figure A.1.1). However, there are no such differences in children between the ages of three and six. Depending on age and federal state, care rates are often over 90% with only minor differences between federal states (Federal Statistical Office 2022a).

Both public and private providers offer care places, which are initially financed by municipalities. Only a small part of the actual costs is passed on to parents by the respective municipalities, so that the places are heavily subsidised overall. Each municipality has its own regulations for determining parental contributions. However, systematic differences in the determination of contributions between municipalities in East and West Germany cannot be identified.

The 1990s were characterised by strong economic uncertainty in East Germany. Since about 2010, however, living conditions in East and West Germany have become more similar. This is reflected in a stronger growth in productivity in East relative to West Germany. Despite the fact that the unemployment rate in East Germany was nearly double that of West Germany until the mid-2000s, there has been a marked convergence since that time (Krause 2019). Supporting Information S1: Figure A.1.2 shows a higher unemployment rate for women in East Germany of 7.3% compared to 5.0% in West Germany in 2018. However, the figure also makes it clear that areas within West Germany, such as the Ruhr area in the West, can also be affected by higher unemployment rates.

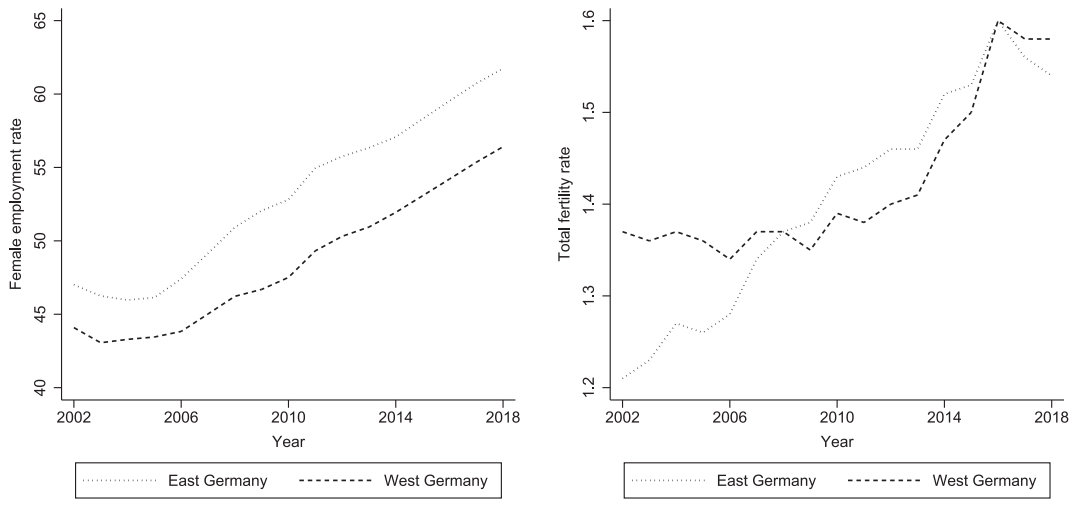


**FIGURE 1** | Fertility and female employment rate in East and West Germany. Own compilation based on Federal Statistical Office 2022b and Bundesagentur Für Arbeit (BA) 2020. The total fertility rate is defined as the sum of age-specific fertility rates (based on women aged 15–49 years). The employment rate indicates the proportion of female employees aged 15 to 64 who are subject to social insurance contributions in the female population of the same age.

## 2.2 | Related Literature

There are several theoretical models that establish a connection between the number of children and the labour supply of mothers (Becker 1985; Mincer 1962). Most of these theoretical models predict a negative relationship between fertility and maternal labour market outcomes. Parenthood leads to specialisation within the household (Becker 1985). Because of the greater biological investment in children (e.g., breastfeeding or pregnancy), women are more willing to invest time in childcare, thereby reducing their labour supply (Becker 1991).

According to theoretical neoclassical models, however, the effect of fertility on maternal labour supply is less clear (see, among others, Rønsen and Sundström (2002) or Leibowitz et al. (1992)). Here, a woman enters the labour market when her full wage exceeds the reservation wage. In addition to the market wage, the full wage also includes the opportunity costs of not working and the expected impact on income due to the decline in human capital caused by not working. The reservation wage, defined as the lowest wage at which the woman is willing to work, is derived from the utility of time at home and the value of home production. On the one hand, motherhood increases the reservation wage because childcare is time-intensive. On the other hand, children increase the demand for market goods. Consequently, the effect of fertility on mothers' employment can be either positive or negative. Because the care of small children is difficult to substitute, the negative effects predominate in the child's first years of life. Therefore, the effect of fertility on maternal labour market outcomes would strongly depend on the earning potential and the preferences regarding paid work. Furthermore, the impacts may be heterogeneous depending on the institutional environment (e.g., the extent of publicly provided childcare), people's social norms (attitudes towards working mothers) and economic situation (number of jobs that allow work–family balance) in the surrounding area.

The difference in maternal employment between West and East Germany has been widely discussed in the economic literature over the last 30 years. In spite of similar family policies for reunified Germany as a whole since 1990, mothers in East Germany show different behaviour with respect to employment and the usage of childcare. These differences can also be explained by cultural values and different family models (Pfau-Effinger 2012). Previous studies find that people who grew up in the GDR developed different social norms regarding the role of mothers in the labour market (Campa and Serafinelli 2019; Lippmann et al. 2020; Bauernschuster and Rainer 2012).

Lippmann et al. (2020) demonstrate that the so-called "male breadwinner norm" has been altered by the institutions of the GDR. According to their results, unlike their West German counterparts, women in East Germany can have a higher income than their husbands without increasing their risk of divorce or being forced to compensate by working more hours at home or leaving the labour market. The role of children has been investigated by Jessen (2022) and Collischon et al. (2020). Their findings suggest that child penalties on labour market outcomes are significantly greater for women in West Germany than in East Germany, which they attribute to different social norms.

Socialist systems also affected preferences regarding social policy (Alesina and Fuchs-Schündeln 2007). This could also explain why, despite the GDR's full integration into the FRG's economic and political system, differences between East and West in the provision of public childcare and prevalence of public employment remained (Rosenfeld et al. 2004). Empirical studies (Baker et al. 2008; Chhaochharia et al. 2021; Huber and Rolvering 2023) show that the level of motherhood penalties is strongly influenced by the childcare provision. Therefore, it can be assumed that institutional framework conditions have contributed to the higher employment of mothers in the East. These studies, however, have not considered the heterogeneous effect of the number of children. It is unclear whether the differences between East and West only arise when entering motherhood, or whether these differences become more pronounced with each additional child.

However, the causal effect of fertility on maternal labour market outcomes is difficult to determine because both may be determined by similar factors. For example, family-oriented women could self-select into motherhood, whereas career-oriented women might have preferences for a smaller number of children (Hakim 2003). In addition, there are other unobserved characteristics, such as preference for leisure time, that influence women's fertility and work decisions. Any estimate that does not control for all these variables would over- or underestimate the true effect. Following Rosenzweig and Wolpin (1980), many studies have therefore used twin births as a natural experiment to examine the causal effect of the number of children on maternal labour market outcomes. This can be considered a good instrument because it increases the desired number of children exogenously and is not correlated with personal characteristics.[4] Different authors have challenged this approach recently by arguing against the exogeneity of the instrument due to possible correlation with mother's health (Farbmacher et al. 2018; Bhalotra and Clarke 2019). Various studies on industrialised countries with this exogenous variation find a negative relationship between fertility and female labour supply (see Supporting Information S1: Table A.1.1).

Most studies refer to only one country. A notable exception is the study by Baranowska-Rataj and Matysiak (2016), which investigates the effect of fertility on female employment for various European country groups. The magnitude of the effect varies between these groups: Only in Nordic countries and Central and Eastern Europe (CEE), no statistically significant effect can be determined. The strongest effect was found in Anglo-Saxon and southern European countries, followed by western-French and other Western countries.

The twin instrument can only examine the question of how an additional child affects women's labour market outcomes (the intensive fertility margin), but not the effect of entering motherhood (the extensive fertility margin). Lundborg et al. (2017) use a dataset of IVF-treated women in Denmark to estimate the causal relationship of childbearing on labour market outcomes. The paper applies an instrumental variable approach by using the success at the first in-vitro-fertilisation (IVF) treatment of childless women with embryo implants as their instrument. Their estimates suggest that the effects are substantially stronger at the extensive margin compared to the intensive margin.

## 3 | Data and Descriptive Statistics

Our analysis of the effects of family size on maternal labour market outcomes in Germany is based on several data sources. The main source of information is individual tax return data for 2018. This data is supplemented by regional information on the availability of childcare services as well as information on regional gender norms. We use the full sample of the German income tax statistics (Federal Statistical Office 2019b) which collects data from the universe of all income taxpayers in Germany. The cross-sectional dataset contains information on about 29.3 million taxpayers who filed a tax return. In this dataset, more than 800 variables are included on different kinds of (taxable) incomes, allowances used, and other information needed to compute the final tax liability of each taxpayer.

The use of administrative tax data contrasts with most empirical studies examining the causal relationship between fertility and maternal labour market outcomes, which use either survey or census data. Using tax data offers at least two advantages. First, tax data provides highly reliable and detailed figures on different incomes, which is not always the case with other data.[5] Furthermore, survey data may exhibit bias, either by misreporting actual income or failing to report it at all. This can particularly affect the income level reported by women in household surveys. Using Swiss data combining survey and administrative information, Roth and Slotwinski (2020) show that women's incomes are systematically underreported when their actual income exceeds that of their male partner. Second, the large number of observations allows us to conduct our analyses on heterogeneous subgroups of the population. Our identification strategy ideally needs a large number of observations with twin births—a relatively rare event. Compared to census or survey data, which cover only a small portion of the total population, the 29.3 million individual tax returns we use represent a much broader database.[6]

However, the use of administrative tax data also comes with disadvantages. One issue is the external validity of our results. Households that are not obliged to file a tax return and do not do so voluntarily are not included in our data. In about 80% of corresponding cases, these are unmarried taxpayers with low incomes who only pay payroll taxes (Federal Ministry of Finance 2024). In addition, socioeconomic variables that are not required for the calculation of taxes, such as education, occupation and hours worked, are only available to a limited extent in the data.

The present study focuses on married couples who are jointly assessed for income tax purposes and have at least one child.[7] Our dataset is restricted to women who are between the ages of 20 and 45 and who were at least 18 years old at the time of their first birth. The oldest child has to be younger than 13 years. The age limit of 13 years was chosen because previous literature shows that the effect of an additional child on maternal labour supply vanishes when the child turns 13 (Bronars and Grogger 1994; J. Angrist and Evans 1998). This restriction also means that we consider a period with similar fertility rates in East and West Germany. We further exclude families with foster and step children, grandchildren and parents who receive or pay maintenance payments.

The identification of twins is based on the year of their birth.[8] We observe 68,278 twins, which corresponds to 1.94% of twins in our data.[9] Our total sample size is approximately 2.07 million observations, with 1.78 million from West Germany and 0.29 million from East Germany (see Table 1). Of these, more than 1.21 million have at least two children and around 225,000 have at least three.

The employment rate of mothers in East and West Germany varies within the whole sample, with 81.0% in East Germany and 65.5% in West Germany (Table 1). A woman is considered to be employed if her labour income is equal to or greater than zero. We define labour income as income from employment.[10] In any case, the participation rate is lower than that of fathers, which is around 90% for the full sample. This rate decreases slightly with the number of children.

The average labour income for mothers in West Germany is only €13,935 per year. In contrast, mothers in East Germany have an average labour income of €19,819.[11] However, these numbers are strongly driven by mothers who are not in employment. When considering only those with positive incomes, the average labour income is €24.456 in East and €21.271 in West Germany. Differences can also be observed in the distribution of positive labour income (see Figure 2). Although both distributions are right-skewed, the mode of the distribution in East Germany lies at a higher income level compared to the West. Moreover, the distribution in the East is more dispersed and shifts further to the right, indicating a greater share of women with higher labour incomes. In contrast, the West German distribution is more concentrated at lower income levels.

The average age of mothers at first birth is 28.77 years in East Germany and 29.35 years in West Germany. On average, mothers with at least two children had their first birth at the age of 28 and 28.64, respectively.[12] Information regarding the regional availability of public childcare facilities and social norms that are not included in tax data can be assigned to individual taxpayers based on their place of residence. Our data on childcare availability includes care rates for all 294 German counties and the three city states of Berlin, Bremen, and Hamburg for the year 2018 (Federal Statistical Office 2019a). The data allows us to distinguish between children in part-time and full-time care (Supporting Information S1: Figure A.1.1). We use the ifo Survey on Gender Norms to measure regional gender norms. This representative opinion survey, which was conducted in Germany in 2020, contains beliefs concerning gender norms and parental labour market behaviour. The dataset also includes regional identifiers ("Raumordnungsregion"). The gender norms employed in our analysis are derived from responses to two attitudinal questions concerning the role of women with young children in the labour market: "*A toddler will certainly suffer if his or her mother works.*" and "*Mothers whose children are not yet in school should reduce their working hours.*" These questions relate to attitudes towards the extent to which caring for young children is the primary responsibility of mothers. The first statement refers to the extensive margin of maternal labour supply, whereas the second refers to the intensive margin. There are major differences between East and West German regions (Supporting Information S1: Figure

**TABLE 1** | Descriptive statistics from tax data.

| | At least one child | | At least two children | | At least three children | |
|---|---|---|---|---|---|---|
| **Variables** | **East** | **West** | **East** | **West** | **East** | **West** |
| Number of children | 1.678 | 1.720 | 2.186 | 2.217 | 3.150 | 3.154 |
| | (0.680) | (0.706) | (0.455) | (0.487) | (0.418) | (0.424) |
| Age of youngest child (years) | 3.929 | 3.542 | 3.539 | 3.497 | 2.473 | 2.592 |
| | (3.131) | (3.038) | (2.700) | (2.770) | (2.157) | (2.235) |
| Age of mother at first birth (years) | 28.770 | 29.350 | 28.000 | 28.640 | 26.850 | 27.210 |
| | (4.122) | (4.336) | (3.882) | (4.132) | (3.938) | (4.142) |
| Family has twins at 1st birth | 0.023 | 0.022 | — | — | — | — |
| | (0.150) | (0.147) | | | | |
| Family has twins at 2nd birth | — | — | 0.015 | 0.016 | — | — |
| | | | (0.121) | (0.125) | | |
| Family has twins at 3rd birth | — | — | — | — | 0.013 | 0.013 |
| | | | | | (0.111) | (0.112) |
| Labour income of mother (in €) | 19,819 | 13,935 | 19,497 | 12,672 | 14,545 | 8507 |
| | (19,043) | (18,672) | (19,111) | (18,346) | (18,579) | (16,727) |
| Income from labour and business activity of mother (in €) | 21,740 | 15,550 | 21,590 | 14,527 | 16,967 | 10,771 |
| | (42,084) | (49,009) | (24,952) | (57,825) | (26,932) | (48,839) |
| Mother is employed | 0.810 | 0.655 | 0.797 | 0.615 | 0.665 | 0.455 |
| | (0.392) | (0.475) | (0.402) | (0.487) | (0.472) | (0.498) |
| Mother has income from business activity | 0.081 | 0.076 | 0.087 | 0.084 | 0.099 | 0.091 |
| | (0.273) | (0.264) | (0.281) | (0.277) | (0.299) | (0.287) |
| Income from business activity of mother (in €) | 1921 | 1616 | 2093 | 1855 | 2422 | 2264 |
| | (38,330) | (45,458) | (17,993) | (54,879) | (20,814) | (45,982) |
| Income from labour and business activity of father (in €) | 46,552 | 59,015 | 49,938 | 63,530 | 54,659 | 68,693 |
| | (85,155) | (149,089) | (91,134) | (176,516) | (113,429) | (322,654) |
| Age of mother (years) | 35.020 | 35.070 | 35.600 | 35.820 | 35.760 | 35.830 |
| | (4.588) | (4.900) | (4.193) | (4.488) | (4.027) | (4.304) |
| Age of father (years) | 38.100 | 38.120 | 38.600 | 38.860 | 39.010 | 39.130 |
| | (5.990) | (6.134) | (5.435) | (5.616) | (5.342) | (5.460) |
| Family unaffiliated with any religion | 0.870 | 0.515 | 0.850 | 0.496 | 0.794 | 0.525 |
| | (0.337) | (0.500) | (0.357) | (0.500) | (0.405) | (0.499) |
| Father is employed | 0.889 | 0.908 | 0.882 | 0.901 | 0.849 | 0.873 |
| | (0.315) | (0.289) | (0.322) | (0.299) | (0.358) | (0.333) |
| Father has income from business activity | 0.192 | 0.208 | 0.209 | 0.232 | 0.259 | 0.271 |
| | (0.394) | (0.406) | (0.407) | (0.422) | (0.438) | (0.445) |
| Additional income of mother (in €) | 238 | 317 | 263 | 370 | 333 | 489 |
| | (2976) | (3485) | (2930) | (3876) | (2759) | (5294) |
| Observations | 286,193 | 1,783,365 | 163,648 | 1,054,595 | 26,460 | 198,638 |

*Note:* Standard deviations in parentheses. Income from business activity includes income from self-employment, business and agriculture and forestry. Additional income includes income from capital, rental and leasing and other income. Family unaffiliated with any religion is a dummy, which is equal to one if both partners do not belong to any religion and zero otherwise.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).

A.2.3). While in West Germany 16.1% agree with the first statement, it is only 9.6% in East Germany (second statement: 17.6% in West Germany and 10.5% in East Germany). However, there are also major differences within West Germany. In the city states and the northern federal states, approval is lower. The responses to the mirrored attitude questions regarding the role

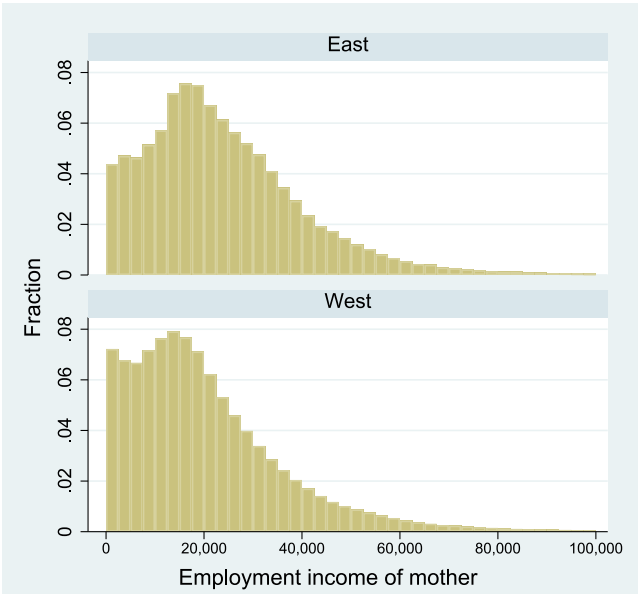


**FIGURE 2** | Distribution of positive female income in East and West Germany. The figure shows the distribution of positive female labour income up to €100.000. For the other samples, see Supporting Information S1: Figures A.4 and A.5 in the Appendix. *Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).

of fathers in childcare indicate low approval ratings, with no significant differences observed between East and West Germany.

## 4 | Estimation Strategy

This study aims to identify the causal effect of family size on maternal employment status and labour income. The OLS estimates may be biased because a greater desire to have children may be associated with preferences regarding labour supply. Therefore, we use twin births as our exogenous variation for family size by applying a two-stage least squares (2SLS) instrumental variable approach.

Using a pooled twin instrument (without taking birth order into account) would lead to an oversampling of women with a higher number of pregnancies and thus also of women with a higher preference for a larger family size (Silles 2015). To adjust our estimation for the connection between the occurrence of twin births and the number of pregnancies, we use three different instruments: twins at first, second and third birth. This allows us to identify the causal effect at different parities.

According to Silles (2015), we estimate the following model for the sample with at least $k$ children:

$$y_i = \beta_0 + \beta_1 n_i^k + \beta_2 X_i + \mu_i, \quad (1)$$

where $y_i$ is the variable of interest (mother's employment status or labour income), $X_i$ is the vector of covariates, $n_i^k$ is the number of children by individual $i$ in the sample of mothers with at least $k$ children ($k \geq 1$, $k \geq 2$, $k \geq 3$) and $\mu_i$ is the error term. Control variables included in $X_i$ are age of mother, age of mother squared, age of mother at first birth and age of youngest child.

The first stage is estimated according to the following equation:

$$n_i^k = \alpha_0 + \alpha_1 t_{ij} + \alpha_2 X_i + \varepsilon_i, \quad (2)$$

where $t_{ij}$ is a dummy whether individual $i$ had twins at $jth$ birth/parity ($j = 1$, $j = 2$, $j = 3$). We use the instrument $t_{i1}$ in the sample with at least one child to measure the effect of the second child, $t_{i2}$ in the sample with at least two children to measure the effect of the third child and $t_{i3}$ in the sample with at least three children to measure the effect of the fourth child. $\varepsilon_i$ is the error term.

For the instrument to be valid, it has to fulfil several conditions. First, the causal effect of the instrument on the endogenous treatment variable must be unequal zero on average. This means that a twin birth has to be correlated with the number of children. This hypothesis can be tested by the results of the first stage estimation (Table 2). Having twins at first, second and third birth is highly correlated with the number of children. The t- and F-statistics demonstrate that the instrument is statistically significant and not weak. The coefficient in the sample with at least one child is larger in East (0.866) than that in West Germany (0.840). In the sample with at least two children, the coefficients are 0.926 in East Germany and 0.898 in West Germany. The coefficients are similar in the sample with at least three children: 0.930 in East and 0.928 in West Germany. Compared to previous studies, the magnitude of our estimates falls within the upper range (see Supporting Information S1: Table A.1.1). However, it should be noted that several of these estimates are not directly comparable to ours because they often include older children. In the long term, families have more time to adjust their fertility after a twin birth, resulting in a smaller effect. Silles (2015), who applies a similar restriction to that in this study, finds a considerably smaller effect of 0.697 for the first birth. The effects of the subsequent births, however, are more similar to our estimates. Baranowska-Rataj and Matysiak (2016) find effects ranging from

**TABLE 2** | First stage results of IV-regression.

| | At least one child | | At least two children | | At least three children | |
|---|---|---|---|---|---|---|
| | East | West | East | West | East | West |
| Twins at first birth | 0.866*** | 0.840*** | | | | |
| | (0.003) | (0.001) | | | | |
| Twins at second birth | | | 0.926*** | 0.898*** | | |
| | | | (0.007) | (0.002) | | |
| Twins at third birth | | | | | 0.930*** | 0.928*** |
| | | | | | (0.017) | (0.006) |
| Observations | 286,193 | 1,783,365 | 163,648 | 1,054,595 | 26,460 | 198,638 |
| F-statistics on excluded instruments | 74,866.714 | 402,994.318 | 19,874.970 | 134,234.468 | 3027.675 | 24,664.252 |
| Kleibergen-Paap rk LM statistic | 6467.057 | 38,356.411 | 2225.625 | 15,290.456 | 305.310 | 2346.040 |
| Anderson-Rubin Waldtest | 5.505 | 503.730 | 7.303 | 224.346 | 4.030 | 28.774 |
| Partial R-squared | 0.106 | 0.093 | 0.078 | 0.074 | 0.066 | 0.066 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

0.61 to 0.80 for the first birth, depending on the country group. These differences may be attributed to the use of survey data in these studies, whereas the present study is based on a full sample of the income tax statistics. Furthermore, our analysis only considers married couples in which both spouses are the legal parents to all children in the household.

Second, the instrument must be randomly assigned. The increased utilisation of fertility treatments, particularly in in-vitro-fertilisation (IVF), has resulted in a rise in twin births in recent years, thereby challenging this assumption. If these treatments are correlated with labour market outcomes, the negative effect of fertility on women's labour market outcomes may be overestimated. However, the study by Braakmann and Wildman (2016) additionally estimates models that control for fertility treatments and finds only minor differences. Specific regulations of German public health insurance might also mitigate this potential bias.[13] The use of IVF is not observable in our data. Table 1 shows a higher probability of twins at first birth. The average values for the second and third births are closer to the natural probability of having twins.[14] This suggests that IVF treatments are used especially for the first birth. However, since the probability of twin births is similar in East and West Germany, we assume that the proportion of IVF treatments is not systematically different in East and West Germany.

In addition, there is a higher incidence of twin births for older women. We address this concern by including age variables for women in our empirical specification. Recent studies show that twinning is also linked to maternal health (Bhalotra and Clarke 2019). Because mothers of twins are healthier, the effect of birth on their labour market outcomes may be smaller. This would lead to an underestimation of the effects. Although the randomness of twin births cannot be tested directly, we examine whether twin births are independent of observable characteristics. The descriptives (Supporting Information S1: Tables A.2–A.4) show that despite small differences, twin births are not systematically associated with variables such as religious affiliation, urbanisation, or employment status of the father. However, women who give birth to twins are on average older, which reflects the positive association between twinning and the woman's age.

Third, the so-called exclusion restriction must be fulfilled. This states that the instrument may only influence the outcome variable indirectly via its direct effect on the endogenous treatment variable. Rosenzweig and Zhang (2009) argue that the twin instrument may violate this exclusion restriction due to their lower birth weight and infant survival rate. Therefore, parents may decide to invest more time resources in twins than in singletons, thus reducing their labour supply. If a twin birth had a direct effect on maternal employment or labour income that does not go through the indirect channel of the number of children, this would be a violation of the exclusion restriction. Rosenzweig and Zhang (2009) solve this potential problem by including birth weight in their empirical specification. Unfortunately, our tax data does not include information on birth weight. However, Spiess and Dunkelberg (2009) find no statistically significant impact of preterm births on mothers' labour supply in Germany. Therefore, this issue may be negligible in the German context. Additionally, there is no indication that this potential issue affects East and West Germany systematically differently.[15] Ultimately, the estimates must be interpreted as a local average treatment effect (LATE) for the group of compliers (J. D. Angrist et al. 1996). For compliers, a twin birth leads to a higher number of children than originally planned.[16]

## 5 | Results

### 5.1 | Main Results

The OLS estimates indicate a statistically significant negative relationship between the number of children and employment status for all samples (Table 3).[17] In the sample with at least one

**TABLE 3** | Effect of an additional child on female employment.

| | OLS-estimation | | IV-estimation | |
|---|---|---|---|---|
| | **East** | **West** | **East** | **West** |
| At least one child | | | | |
| Number of children | −0.092*** | −0.101*** | −0.013** | −0.062*** |
| | (0.002) | (0.001) | (0.005) | (0.003) |
| % Effect | 11.4% | 15.4% | 1.6% | 9.5% |
| Adjusted R-squared | 0.038 | 0.071 | 0.031 | 0.070 |
| Observations | 286,193 | 1,783,365 | 286,193 | 1,783,365 |
| At least two children | | | | |
| Number of children | −0.151*** | −0.135*** | −0.024*** | −0.063*** |
| | (0.003) | (0.001) | (0.009) | (0.004) |
| % Effect | 18.9% | 22.0% | 3.0% | 10.2% |
| Adjusted R-squared | 0.057 | 0.091 | 0.042 | 0.087 |
| Observations | 163,648 | 1,054,595 | 163,648 | 1,054,595 |
| At least three children | | | | |
| Number of children | −0.191*** | −0.118*** | −0.057** | −0.056*** |
| | (0.007) | (0.002) | (0.028) | (0.010) |
| % Effect | 28.7% | 25.9% | 8.5% | 12.2% |
| Adjusted R-squared | 0.070 | 0.086 | 0.057 | 0.084 |
| Observations | 26,460 | 198,638 | 26,460 | 198,638 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

child, for example, the probability of employment for mothers with another child in East Germany decreases by 9.2 pp (11.4%) in East Germany and 10.1 pp (15.4%) in West Germany. Although the coefficients from the IV estimations are substantially smaller, they are all still statistically significant at the 1% or 5% level. Accordingly, having an additional child decreases the probability of being employed in the East by 1.3 pp (1.6%) and in the West by 6.2 pp (9.5%). Similar differences between OLS and IV estimates can be observed in all samples. Consequently, our findings indicate that OLS may overestimate the negative relationship caused by an endogenous fertility decision.[18] Women with low labour market prospects may tend to have more children and mothers who desire more children may also prefer to reduce labour supply.

Furthermore, the estimates for the second parity also show stronger effects in East Germany (6.3 pp or 10.2%) compared to West Germany (2.4 pp or 2.0%). However, at the third parity, the coefficients are very similar at 5.7 pp in West and 5.8 pp in East Germany.[19]

Compared with the findings of other studies, our results confirm the negative impact of the number of children on maternal employment in Western countries (see Supporting Information S1: Table A.1.1). Our results are at the lower end of these various estimates. Silles (2015), for example, finds for the second child an effect of 8.7 pp for the UK. While Baranowska-Rataj and Matysiak (2016) do not find statistically significant results for Central East European countries, this is not the case for us. Our estimation results for East Germany are considerably lower than those for West Germany, but the effects are statistically significant.

Interestingly, we observe smaller differences between the OLS and IV estimates in East Germany for at least one and at least two children. In line with Baranowska-Rataj and Matysiak (2016), we interpret this as evidence that selection effects may be more pronounced in the East than in the West. This pattern can be explained by differing normative and institutional contexts between the two regions. In West Germany, the male-breadwinner model has traditionally shaped maternal employment behaviour, leading many women to withdraw—fully or partially—from the labour market following childbirth (Trappe and Rosenfeld 2000; Bauernschuster and Rainer 2012). As a result, additional self-selection into motherhood has a relatively limited marginal impact on labour market outcomes.

In contrast, East Germany has historically followed a more egalitarian family model, supported by institutions that facilitated high female labour force participation (Fuchs-Schündeln and Masella 2016). In this context, self-selection into higher fertility is likely to reflect stronger unobserved heterogeneity, leading to greater endogeneity bias in OLS estimates.

Within the sample of families with at least three children, the differences between OLS and IV estimates in East and West Germany are substantially smaller. This suggests that the

previously discussed distinctions in labour market participation and caregiving models between the two regions may no longer apply in the same way once a certain number of children is present. At higher levels of fertility, structural constraints and personal preferences may converge, reducing the role of region-specific selection effects.

The estimates in Table 4 also suggest a negative relationship between the number of children and labour income. Even though the coefficients from the OLS specification are larger than those from IV, the differences are smaller than for the employment probability estimates. This could indicate that self-selection into larger families particularly influences the effect on the extensive margin of female labour supply. For example, according to the OLS coefficients, an additional child reduces the labour income of mothers in West Germany by on average €2297 (16.5%), whereas the IV results only imply a reduction on average of €1758 (12.6%).

Looking at the results from the IV specification, we find stronger effects for West Germany than for East Germany. For women with at least one child, the effect is almost twice as large at €1758 (12.6%) in West Germany compared to €907 (4.6%) in East Germany. While the effect for West Germany is reduced at higher parities, the opposite is the case for East Germany. The effect of the second child is already €1386 (4.6%) and increases to €1837 (7.1%) for the fourth child (only significant at the 10% level). Thus, the nominal effect in East Germany is significantly higher at the third parity than in West Germany (€1114 or 13.1%) but has similar percentage effects due to different average incomes.[20] However, the extent to which these results are driven by the extensive margin of labour supply remains unclear. Consequently, additional estimates were made with a sample of women with positive income, with log labour income serving as the dependent variable. It must be borne in mind that the estimates for East and West Germany are only comparable to a limited extent. This is due to the different employment rates in the two regions, which are reflected in the number of observations (see Table 5). The estimates show that the labour income of working mothers with at least one child decreases by 7.2% in the East and 8.6% in the West with an additional child. For the third child, the decline is 4.4% in the East and 7.3% in the West. No statistically significant effects were found for the fourth child. These findings indicate that the effects for the first two parities are still lower in East than in West Germany. However, it appears that the estimates for labour income at the third parity are purely driven by the extensive margin of labour supply.

In short, the second and third child generally has a stronger negative impact on maternal labour market outcomes in West than in East Germany. This means that the differences between East and West Germany in women's employment and labour income, are not solely caused by entering parenthood, but are also exacerbated by additional children. The effect decreases in West Germany with the number of children, whereas the opposite is the case in East Germany. As a result, the effects even out with the fourth child.[21]

**TABLE 4** | Effect of an additional child on female labour income.

| | OLS-estimation | | IV-estimation | |
|---|---|---|---|---|
| | **East** | **West** | **East** | **West** |
| At least one child | | | | |
| Number of children | −1334*** | −2297*** | −907*** | −1758*** |
| | (83) | (31) | (283) | (116) |
| %-Effect | 6.7% | 16.5% | 4.6% | 12.6% |
| Adjusted R-squared | 0.082 | 0.091 | 0.082 | 0.091 |
| Observations | 286,193 | 1,783,365 | 286,193 | 1,783,365 |
| At least two children | | | | |
| Number of children | −3558*** | −2518*** | −1386*** | −1657*** |
| | (103) | (38) | (418) | (169) |
| %-Effect | 18.2% | 19.9% | 7.1% | 13.1% |
| Adjusted R-squared | 0.112 | 0.100 | 0.110 | 0.100 |
| Observations | 163,648 | 1,054,595 | 163,648 | 1,054,595 |
| At least three children | | | | |
| Number of children | −3982*** | −1646*** | −1837* | −1114*** |
| | (230) | (73) | (1012) | (322) |
| %-Effect | 27.4% | 19.3% | 12.6% | 13.1% |
| Adjusted R-squared | 0.109 | 0.087 | 0.107 | 0.087 |
| Observations | 26,460 | 198,638 | 26,460 | 198,638 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

**TABLE 5** | Effect of an additional child on logarithm of female labour income.

| | At least one child | | At least two children | | At least three children | |
|---|---|---|---|---|---|---|
| | East | West | East | West | East | West |
| Number of children | −0.072*** | −0.086*** | −0.044** | −0.073*** | −0.062 | −0.053 |
| | (0.014) | (0.007) | (0.023) | (0.012) | (0.081) | (0.042) |
| Adjusted R-squared | 0.106 | 0.100 | 0.136 | 0.117 | 0.121 | 0.103 |
| Observations | 231,931 | 1,168,306 | 130,472 | 648,064 | 17,584 | 90,299 |

*Note:* The table shows two-stage least squares estimates with log female labour income as the dependent variable. Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

## 5.2 | Subgroup Analysis

### 5.2.1 | Urban and Rural Areas

Next, we aim to identify the factors that may lead to the different results in East and West Germany. First, the effect heterogeneity by degree of urbanisation is investigated. Within the EU, women living in cities have a significantly higher employment rate than women in towns, suburbs or rural areas (Eurostat 2012). In addition, a higher degree of urbanity is associated with higher average educational attainment (European Commission 2024). There are substantial differences in urbanisation and the urban–rural divide between East and West Germany (Krause 2019).

The division into different regional types is based on the regional statistical spatial typology for mobility and transport research (Sigismund 2018). We use the differentiated regional statistical regional types, which divide the regions into four categories: metropolitan urban region, regiopolitan urban region, rural region close to an urban region and peripheral rural region.[22]

Overall, as shown in Supporting Information S1: Table A.5, East Germany has larger proportions of particularly high and low levels of urbanisation.[23] In West Germany, however, unlike in the East, a decline in the level of urbanisation with the number of children can be observed. The regression results for employment probability show significant differences between East and West Germany (see Table 6). For West Germany, negative statistically significant results are found in all groups. The smallest effects is present in metropolitan urban regions, with small differences in parity. In peripheral rural regions, however, the effect increases with parity, with 9.0 pp (20.3%) for the fourth child. However, there are no statistically significant results for peripheral rural regions in East Germany. For metropolitan urban regions, there are only statistically significant results for the fourth child. A similar picture emerges for labour income (see Supporting Information S1: Table A.6). In conclusion, mothers in East and West Germany, especially in peripheral rural areas, face significantly different regional circumstances that influence their employment decision and income.

### 5.2.2 | Women With Toddlers

Next, we will determine the effect on mothers with toddlers. According to theoretical literature, this group is expected to demonstrate a stronger effect. This is because young children are more time-consuming for parents and childcare is difficult to substitute. The analysis of this group can also contribute to a better understanding of the reasons for the different effects in East and West Germany.

Previous studies (Baranowska-Rataj and Matysiak 2016; Fontaine 2017; Aaronson et al. 2021) have identified three primary factors influencing mothers' work–family reconciliation. First, the relationship between fertility and female employment rates may be strongly dependent on regional economic development. For example, this could occur, if a suitable job cannot be found due to a high unemployment rate. Secondly, institutional framework conditions, such as the availability and quality of childcare in the surrounding area, can influence the labour supply of women. Providing public childcare reduces costs for mothers to enter the labour market. In addition, the availability of full-day childcare can enable women to work more hours and thereby increase their labour income. Thirdly, regional social norms regarding maternal employment can also have a significant impact. We hypothesised that high levels of support for gender norms of people in the surrounding area increase the costs for mothers to enter employment or to increase their labour supply.

To determine more precisely which mechanisms influence the difference between East and West Germany, the effects of various indicators are tested (see Table 7). In comparison to our baseline specification, the number of children is additionally interacted with the various indicators. The individual samples are restricted to women with a first, second or third child between the ages of one and three. Because the differences in the area of childcare and social norms arise particularly between East and West Germany, rather than within them, the individual estimates are carried out for the entire country.[24] Following Baranowska-Rataj and Matysiak (2022), we estimate the following 2SLS-models:

$$n_i^k = \lambda_0 + \lambda_1 t_{ij} + \lambda_2 X_i + \lambda_3 indicator_i + \lambda_3 interaction_i + \lambda_4 East_i + \phi_i, \quad (3)$$

$$y_i = \sigma_0 + \sigma_1 n_i^k + \sigma_2 X_i + \sigma_3 indicator_i + \sigma_3 interaction_i + \sigma_4 East_i + \psi_i, \quad (4)$$

where *indicator* represents the regional childcare rate, social norms or female unemployment rate. *Interaction* represents an interaction variable between the indicator and family size. *East* is a binary variable that takes the value 1 if the women lives in East

**TABLE 6** | Effect of an additional child on female employment by degree of urbanisation.

| | Metropolitan urban region | | Regiopolitan urban region | | Rural region close to an urban region | | Peripheral rural region | |
|---|---|---|---|---|---|---|---|---|
| | East | West | East | West | East | West | East | West |
| At least one child | | | | | | | | |
| Number of children | −0.011 | −0.056*** | −0.021 | −0.075*** | −0.038*** | −0.065*** | 0.003 | −0.060*** |
| | (0.008) | (0.004) | (0.013) | (0.006) | (0.014) | (0.006) | (0.011) | (0.007) |
| % Effect | 1.4% | 8.3% | 2.5% | 11.4% | 4.5% | 10.2% | 0.4% | 9.4% |
| Adjusted R-squared | 0.031 | 0.072 | 0.036 | 0.073 | 0.036 | 0.071 | 0.026 | 0.062 |
| Observations | 135,676 | 775,336 | 42,691 | 374,439 | 41,069 | 357,892 | 62,994 | 274,582 |
| At least two children | | | | | | | | |
| Number of children | −0.017 | −0.055*** | −0.038* | −0.076*** | −0.076*** | −0.056*** | −0.002 | −0.074*** |
| | (0.013) | (0.006) | (0.023) | (0.009) | (0.023) | (0.009) | (0.018) | (0.011) |
| % Effect | 2.2% | 8.7% | 4.7% | 12.4% | 9.1% | 9.3% | 0.2% | 12.5% |
| Adjusted R-squared | 0.039 | 0.091 | 0.050 | 0.092 | 0.061 | 0.086 | 0.036 | 0.076 |
| Observations | 77,083 | 443,637 | 24,441 | 222,532 | 23,413 | 217,878 | 36,511 | 169,898 |
| At least three children | | | | | | | | |
| Number of children | −0.068* | −0.046*** | −0.028 | −0.049** | −0.082 | −0.051** | −0.049 | −0.090*** |
| | (0.041) | (0.016) | (0.074) | (0.022) | (0.079) | (0.023) | (0.058) | (0.024) |
| % Effect | 10.3% | 9.9% | 4.2% | 10.8% | 12.4% | 11.6% | 7.1% | 20.3% |
| Adjusted R-squared | 0.054 | 0.090 | 0.065 | 0.087 | 0.064 | 0.084 | 0.061 | 0.067 |
| Observations | 12,855 | 78,993 | 3875 | 42,749 | 3496 | 43,151 | 5898 | 33,626 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

**TABLE 7** | Indicators for potential mechanisms by outcome variable.

| Mechanism | Employment | Labour income |
|---|---|---|
| 1. Regional economic situation | Female unemployment rate (Supporting Information S1: Figure A.1.2) | |
| 2. Regional institutional framework conditions | Share of children under 3 years in: | |
| | day care centres (Supporting Information S1: Figure A.1.1) | All-day care centres (Supporting Information S1: Figure A.1.1) |
| 3. Regional social norms | Proportion strongly agreeing to the statement: | |
| | "*A toddler will certainly suffer if his or her mother works*." (Supporting Information S1: Figure A.2.3) | "*Mothers whose children are not yet in school should reduce their working hours*." (Supporting Information S1: Figure A.2.3) |

*Source:* Own compilation.

Germany and 0 in the West. $\phi$ and $\psi$ describes the respective error terms. All other variables are as in Equations (1) and (2).

The subsamples show lower average labour income and employment rates in comparison to the main analysis (see Supporting Information S1: Table A.7). For instance, the mean employment rate among households with at least one child is 61.9% (76.5% in the East and 60.0% in the West) and the mean labour income is €11,789 (€16,388 in the East and €11,159 in the West).

The overall effect size on employment status is larger than that in the main analysis (Table 8). For instance, among those with at least two children, the impact of having an additional child is −9.6 pp (17%). The coefficient for East Germany is positive and statistically significant. In the specifications with the childcare rate the coefficient for the number of children increases from −9.6 pp to −15.7 pp. However, the childcare coefficients are positive and highly significant. This suggests that a higher childcare rate may mitigate the negative influence of the number of children on employment.[25] By contrast, the coefficient for East Germany falls sharply, from 22.4 pp to 5.2 pp. This finding indicates that the observed differences between West and East Germany are primarily attributable to the different levels of childcare provision.

Column (7) shows the estimations that take regional gender norms into account. The coefficient for the interaction term of

**TABLE 8** | Effect of an additional child on female employment for women with a toddler.

| | At least one child | | | | At least two children | | | | At least three children | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Number of children | −0.087*** | −0.164*** | −0.057*** | −0.116*** | −0.096*** | −0.157*** | −0.045** | −0.131*** | −0.077*** | −0.063 | −0.146*** | −0.101*** |
| | (0.004) | (0.017) | (0.017) | (0.013) | (0.006) | (0.020) | (0.020) | (0.015) | (0.013) | (0.045) | (0.045) | (0.032) |
| East Germany | 0.154*** | 0.021*** | 0.127*** | 0.155*** | 0.224*** | 0.052*** | 0.193*** | 0.226*** | 0.257*** | 0.087*** | 0.225*** | 0.263*** |
| | (0.002) | (0.003) | (0.002) | (0.002) | (0.002) | (0.003) | (0.002) | (0.002) | (0.004) | (0.007) | (0.005) | (0.005) |
| Share day care | | 0.042*** | | | | 0.049*** | | | | 0.122** | | |
| | | (0.009) | | | | (0.017) | | | | (0.058) | | |
| Number children # Share day care | | 0.032*** | | | | 0.026*** | | | | −0.007 | | |
| | | (0.007) | | | | (0.008) | | | | (0.019) | | |
| Toddler suffer | | | −0.132 | | | | 0.243 | | | | −1.814** | |
| | | | (0.143) | | | | (0.258) | | | | (0.849) | |
| Number children # Toddler suffer | | | −0.200* | | | | −0.334*** | | | | 0.444 | |
| | | | (0.109) | | | | (0.123) | | | | (0.275) | |
| Fem. unemp. | | | | −0.008** | | | | −0.017*** | | | | −0.019 |
| | | | | (0.003) | | | | (0.006) | | | | (0.020) |
| Number children # Fem. unemp. | | | | 0.006** | | | | 0.008*** | | | | 0.005 |
| | | | | (0.003) | | | | (0.003) | | | | (0.006) |
| Average employment rate | 0.619 | | | | 0.579 | | | | 0.416 | | | |
| Adjusted R-squared | 0.091 | 0.097 | 0.093 | 0.091 | 0.109 | 0.117 | 0.110 | 0.108 | 0.103 | 0.111 | 0.104 | 0.103 |
| Observations | 514,553 | 514,553 | 514,553 | 514,553 | 422,356 | 422,356 | 422,356 | 422,356 | 101,556 | 101,556 | 101,556 | 101,556 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child. The samples include mothers with their first (sample with at least one child), second (sample with at least two children) or third child (sample with at least three children) aged 1–3 years.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

gender norms with the number of children is statistically significant and negative, indicating that high conformity to these norms has a negative impact on employment.[26] Despite the large effect size of social norms, the coefficient for East Germany is only slightly reduced to 19.3 pp. From this, we can conclude that social norms influence the effect size in East and West Germany. However, they can only explain the differences between the two regions to a limited extent.

The specification with the female unemployment rate leads to only minor changes in the coefficient for East Germany. This suggests that the regional economic situation does not provide a sufficient explanation for the different effects in East and West Germany.

The number of children significantly impacts the labour income of mothers with young children. The sample with at least one child, for example, shows an effect of -€2981 (25%). Residing in East Germany has a statistically significant positive effect on labour income. The effect is reduced by more than 100% in the specification with the all-day care rate to - €389. This pattern is also evident, albeit to a lesser extent, for the other parities. In the specification with the gender norm, the coefficient of the number of children is reduced to -€1.761.[27] Overall, the effect of East German residence is reduced by 17–31% when social norms are taken into account. Furthermore, the third specification indicates that variations in the female unemployment rate can only partially explain the differences between East and West Germany. In the sample with at least one child, the East German effect is slightly reduced to €4284.

Our results suggest that institutional conditions, specifically the availability of public childcare, are the primary explanation for the differing effects of fertility on employment status and labour income in East and West Germany. These results align with literature on the significant impact of public childcare expansion on maternal employment (Baker et al. 2008; Chhaochharia et al. 2021).

Social norms in the immediate environment have a smaller influence on maternal employment, but somewhat stronger effects on labour income. The results are in line with the literature on the relationship between gender norms and parenthood. Baranowska-Rataj and Matysiak (2022) show that attitudes in a country towards the role of men in a family can influence the division of labour in the household (Table 9).

### 5.3 | Robustness Checks

According to J. D. Angrist and Imbens (1995), instrumental variable regressions should be interpreted as a local average treatment effect (LATE). For a better understanding of our estimates, the sample can be divided into two groups: the compliers and always takers. The subgroup of compliers consists of women who have another child because of the instrument. In contrast, the always takers have another child regardless of the instrument. In an ideal scenario, an analysis would exclusively encompass the groups of compliers. However, information on preference regarding the number of children is not available in the dataset. The following picture emerges when looking at mothers aged 40–54 in Germany: 32% have one child, 47% two children, 15% three children and 6% four or more children (Federal Statistical Office 2019a). From this we can conclude that the proportion of always takers is highest in the subgroup with at least one child. If the proportion of compliers differs systematically between East and West Germany, this could distort our results. The higher coefficients in the first stage for the first two parities in East Germany might indicate a slightly larger proportion of compliers in East Germany. Therefore, we estimate a model based on Zhang (2017) with samples with the same frequency of deliveries (see Table 10). To this end, we carry out simple OLS estimates with twins as the independent variable.[28] Thereby, the group of always takers without twin births is excluded from the sample. Estimates of the impact of having a second child on the labour supply are much smaller (see Table 11). The coefficient for East Germany is statistically insignificant. Results from the other samples show statistically significant negative effects, which are slightly smaller. However, the differences between East and West Germany remain. Overall, these findings suggest that our results are not biased by different shares of compliers in East and West Germany.[29]

In the subgroup with at least two children, the gender of the first two children could also influence the fertility decision. This hypothesis could be tested with a combined same-sex and twins instrument. However, our dataset does not include the gender of the child. Previous studies (Fontaine 2017; Silles 2015) have shown only minor differences between these two instruments.

Our results may be related to the different mating patterns in East and West Germany. Grave and Schmidt (2012) show that assortative mating is more common in East than in West Germany. A higher level of homogamy could lead to less specialisation in the household and thus positively influence the woman's labour supply. To test this hypothesis, we estimate models where we control for various characteristics of the husband and the couple's religious affiliation (see Supporting Information S1: Table A.10 and A.11). The coefficients are of a similar magnitude and the differences between East and West Germany remain. It can therefore be assumed that different mating patterns in East and West Germany do not distort our results.[30,31]

## 6 | Conclusion and Discussions

This paper examines the causal effect of the number of children on maternal employment and labour income in East and West Germany using administrative tax data. Multiple births are used as the exogenous variation for the number of children.

Our findings reveal a consistently negative impact of additional children on maternal labour market outcomes, aligning with prior literature. However, the analysis uncovers substantial heterogeneity in the magnitude of these effects. OLS estimates tend to overstate the negative association, suggesting that women with weaker labour market attachment are more likely to have larger families. More importantly, the causal estimates

**TABLE 9** | Effect of an additional child on female labour income for women with a toddler.

| | At least one child | | | | At least two children | | | | At least three children | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Number of children | −2981*** | −4919*** | −1761** | −3772*** | −2667*** | −3088*** | −1475 | −3649*** | −1693*** | 179 | −4351*** | −2723*** |
| | (163) | (626) | (695) | (456) | (225) | (713) | (903) | (478) | (404) | (1505) | (1427) | (964) |
| East Germany | 4727*** | −389*** | 3284*** | 4385*** | 6929*** | 756*** | 5281*** | 6500*** | 6915*** | 1951*** | 5708*** | 6566*** |
| | (76) | (128) | (90) | (78) | (77) | (140) | (92) | (80) | (166) | (286) | (194) | (170) |
| All-day care | | 817*** | | | | 187 | | | | −816 | | |
| | | (264) | | | | (311) | | | | (664) | | |
| Number children #All-day care | | 2140*** | | | | 3299*** | | | | 5495*** | | |
| | | (347) | | | | (650) | | | | (2049) | | |
| ReduceLFP | | | −10,612** | | | | −7485 | | | | −64,454*** | |
| | | | (5204) | | | | (9946) | | | | (24,239) | |
| Number children #ReduceLFP | | | −7316* | | | | −7038 | | | | 15,716** | |
| | | | (3950) | | | | (4755) | | | | (7842) | |
| Fem. unemp. | | | | −21 | | | | −200 | | | | −528 |
| | | | | (120) | | | | (209) | | | | (676) |
| Number children #Fem. unemp. | | | | 163* | | | | 212** | | | | 234 |
| | | | | (92) | | | | (100) | | | | (220) |
| Average labour income | 11,789 | | | | 11,562 | | | | 7917 | | | |
| Adjusted R-squared | 0.127 | 0.133 | 0.129 | 0.127 | 0.125 | 0.133 | 0.127 | 0.125 | 0.095 | 0.101 | 0.097 | 0.095 |
| Observations | 514,553 | 514,553 | 514,553 | 514,553 | 422,356 | 422,356 | 422,356 | 422,356 | 101,556 | 101,556 | 101,556 | 101,556 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child. The samples include mothers with their first (sample with at least one child), second (sample with at least two children) or third child (sample with at least three children) aged 1–3 years.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

**TABLE 10** | Description of subsamples for efficient instrument by Zhang.

| | Sub-sample A | Sub-sample B | Sub-sample C |
|---|---|---|---|
| Mothers of non-twins | (1) | (1)(2) | (1)(2)(3) |
| Versus | | | |
| Mothers of twins | ●1 ●1 | (1) ●2 ●2 | (1)(2) ●3 ●3 |
| Number of children | 1 versus 2 | 2 versus 3 | 3 versus 4 |

*Note:* Own compilation based on Zhang (2017). The numbers correspond to the birth order. ● are twins and ( ) singletons.

**TABLE 11** | Effect of an additional child on female employment (Efficient instrument by Zhang).

| | Sub-sample A | | Sub-sample B | | Sub-sample C | |
|---|---|---|---|---|---|---|
| | East | West | East | West | East | West |
| Twins at first birth | 0.005 | −0.038*** | | | | |
| | (0.005) | (0.003) | | | | |
| Twins at second birth | | | −0.026*** | −0.072*** | | |
| | | | (0.008) | (0.004) | | |
| Twins at third birth | | | | | −0.048* | −0.058*** |
| | | | | | (0.027) | (0.010) |
| %-Effect | 0.5% | 5.4% | 3.1% | 11.1% | 7.0% | 12.5% |
| Adjusted R-squared | 0.025 | 0.033 | 0.032 | 0.065 | 0.044 | 0.069 |
| Observations | 128,093 | 761,229 | 133,885 | 838,931 | 20,196 | 153,106 |

*Note:* Robust standard errors in parentheses. Control variables include age of mother, age of mother squared, age of mother at first birth and age of youngest child.
*Source:* Own computations based on German income tax statistics 2018 (Federal Statistical Office 2019b).
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

demonstrate that the negative effect intensifies with each additional child, particularly in West Germany. The second and third child lead to significantly stronger reductions in maternal employment and income in the West compared to the East. These regional differences narrow only with the fourth child, at which point the effects converge. This pattern indicates that the impact of fertility on female labour outcomes is not uniform but highly dependent on both birth order and regional context. Thus, the East-West gap in maternal labour supply is not driven solely by the transition into motherhood but increasingly reinforced with each additional child.

A subgroup analysis for mothers with toddlers reveals that these differences in East and West Germany can largely be explained by different family policies, particularly the provision of publicly funded childcare. Our results also indicate that social norms in the regional environment have an influence on maternal labour market outcomes, albeit to a lesser extent. This suggests that violating prevailing social norms leads to costs that prevent women from entering the workforce.

The question of how strongly personal social norms influence this causal relationship remains unanswered. Boelmann et al. (2021) show that East German mothers who live in the West continue to have a high level of employment. West German women who live in the East, on the other hand, adjust their labour supply to that of East Germans. Therefore, personal attitudes can reinforce the effect of regional egalitarian gender norms on mother' labour supply. The question also arises about the extent to which mothers know about the attitudes of people around them. Cortés et al. (2022) show that there are misperceptions concerning social norms in the immediate environment, leading to an overestimation of the degree of gender conservativeness. These information gaps could particularly influence the behaviour of mothers in the more attitude-heterogeneous West Germany.

However, family policy can also shape social norms in a country (Ellingsæter et al. 2017; Zoch and Schober 2018). For example, Zoch and Schober (2018) show that the expansion of childcare in West Germany led to more egalitarian gender ideologies among mothers in the short term. The existence of short-term effects indicates that the expansion of childcare must be a central component of a family policy that simultaneously promotes fertility and female employment. This means that despite different social norms, family policy measures, along with tax incentives (Selin 2014), remain good instruments for increasing female employment.


### Acknowledgements

We thank conference participants at IIPF 2020, ESPE 2021 and VfS Annual Conference 2020 for their comments and suggestions. We also thank the participants of the internal seminars at Fraunhofer FIT for their feedback. Additionally, we especially thank Marianna Schaubert for her invaluable help and support. Open Access funding enabled and organized by Projekt DEAL.


### Conflicts of Interest

The authors declare no conflicts of interest.

## Endnotes

[1] For example Frenette (2011) for Canada, Silles (2015) for the UK, Vere (2011) for the US, Fontaine (2017) for France, Moschion (2013) for Australia. Karbownik and Myck (2016) find for Poland negative effects for women with at least one child, but not statistically significant effects for women with at least two children.

[2] East Germany is defined as the territory of the former GDR including Berlin. Since the greater Berlin area represents an important labour market within the East, it is also included. However, we estimate regressions for East Germany excluding Berlin as a robustness check.

[3] For more information see Klinkhammer et al. (2018).

[4] This does not apply in all cases: Firstly, the probability of multiple births increases with the age of the mother. Secondly, women who have undergone a in-vitro-fertilisation (IVF) treatment are more likely to have twins.

[5] The German census, for example, completely lacks information on individual gross earnings. It only states brackets for net income after taxes and social transfers.

[6] German census, for example, only includes a one percent sample of the whole population and an absolute number of about 0.5 million representative households.

[7] In Germany, married couples are usually assessed jointly for the income tax. In principle, German tax law also allows both spouses to make separate assessments. This can be particularly advantageous for the couple if their income is subject to a double taxation agreement. Based on information from the Federal Ministry of Finance (Federal Ministry of Finance (2024)), it can be determined that this option was actually used by less than 4 percent of married couples who filed a tax return in 2018.

[8] Vere (2011) shows that more precise information on the day of birth does not significantly change the results.

[9] Compared to official statistics, this is slightly larger than the proportion of twins to all births in Germany between the years 2000–2018, which is about 1.65 percent (Federal Statistical Office 2024c, 2024d).

[10] Labour income does not include income that is exempt from tax, such as some holiday and night shift bonuses, as well as income from marginal employment.

[11] In our analysis, we restrict the sample to cases where the woman's income is not negative. However, when analysing the effect on income, we include people with zero income. Considering only mothers with positive income would lead to a selection bias and thus make a causal interpretation impossible (J. D. Angrist and Pischke 2009).

[12] These figures are also in line with official statistics, according to which the age at the first birth increased from 28.8 years in 2009 up to 30 years in 2018 (Federal Statistical Office 2024a).

[13] The twin instrument may be biased if higher income, necessary to cover the costs of the treatment, increases the probability of usage. For Germany, however, this should be only a minor problem since the costs of the treatment is in many cases (at least partially) covered by the mandatory public health insurance. In 2018, approximately 21,000 children were born through artificial insemination (D. I. et al. 2020).

[14] Hellin's law describes the probability of twins as 1:89 (Fellman and Eriksson 2009).

[15] J. D. Angrist and Pischke (2009) list a fourth assumption: monotonicity. This states that there are no defiers in the sample. In our case, the condition would not be met if a twin birth would make another child less likely. Since we can rule out this scenario, we consider this fourth assumption to be fulfilled.

[16] This is discussed in Section 5.3.

[17] To determine the effect on the probability of being employed, we estimate linear probability models. This follows the literature on this topic and allows us to directly compare our results with other studies. Since employment is a common event, linear probability models should perform well (Karbownik and Myck 2016).

[18] When comparing the results, it must be taken into account that the OLS estimates represent an ATE (average treatment effect), while the IV estimates represent a LATE.

[19] We also carried out estimates of the employment probability for East Germany without Berlin. The effects are similar, but slightly stronger. However, the differences compared to West Germany remain.

[20] The results for labour income for East Germany are strongly driven by Berlin. Estimates without Berlin show significantly higher coefficients for the first two parities. However, no statistically significant effects can be determined for the fourth child.

[21] Farbmacher et al. (2018) show that the classic twin instrument we use underestimates the negative effect of number of children on income. Therefore, our results might represent a lower bound. Our results show smaller effects than those in the literature (Silles 2015; Karbownik and Myck 2016). Lundborg et al. (2017), however, finds an effect of 5 percent impact of IVF-treated women at the intensive margin.

[22] A metropolitan urban region is a city with a threshold of approximately 500,000 inhabitants or a city region with a threshold of approximately 1 million inhabitants. Regiopolitan urban region are the remaining major cities outside of these metropolitan urban regions. Among rural regions, rural region close to an urban region are defined as regions with a travel time of less than 45 min to the nearest large city or a population potential of more than 300,000 inhabitants. The remaining regions are classified as peripheral rural regions. Supporting Information S1: Figure A.6 shows how these regional types are distributed across Germany.

[23] 99.8 percent of the observations in our dataset could be assigned to a regional type.

[24] The estimates for East and West Germany for this sample are presented in Supporting Information S1: Table A.8.

[25] The variable was transformed so that the coefficients represent the effect of the minimum rate of 13.6 percent. For example, a childcare rate of 30 percent would lead to an effect of an additional child of 10.0 pp.

[26] The coefficients show a hypothetical 100 percent agreement with the statement. For example, 30 percent approval would lead to an effect of an additional child of around −14.5 pp.

[27] This means in the hypothetical case of 0 percent agreement with the norm, the effect of fertility on maternal labour income would be over 40 percent lower.

[28] The exact procedure is described in Supporting Information S1: Subsection A.4.4.

[29] The effects for labour income (see Supporting Information S1: A.9) also confirm our results.

[30] Another reason could be the different educational levels of mothers in East and West Germany. Women with a higher level of education have higher opportunity costs if they stay at home to look after their children. If East German women have a higher level of education than their West German counterparts, this could explain the stronger effect in West Germany. Unfortunately, our dataset does not contain any information on the woman's level of education. Federal Statistical Office (2024b) shows that among 20–44 year-old mothers, the proportion with a polytechnic degree is approximately the same in both parts of the country.

[31] We also carried out an estimate based on the age of the first birth (see Supporting Information S1: Table A.12 and A.13). The age at first birth can affect, among other things, the woman's work experience and education and thus influence whether and what kind of employment can be found. Our estimates show, that the effect in East

Germany is driven by women who gave birth for the first time before the age of 30. There are no statistically significant results for those over 30. In West Germany, however, there are statistically significant effects for women who gave birth for the first time after the age of 30. Since the differences between East and West continue to exist, it can be assumed that differences regarding age at first birth does not confound our results.

## References


Aaronson, D., R. Dehejia, A. Jordan, C. Pop-Eleches, C. Samii, and K. Schulze. 2021. "The Effect of Fertility on Mothers' Labor Supply Over the Last Two Centuries." *Economic Journal* 131, no. 633: 1–32. https://doi.org/10.1093/ej/ueaa100.

Agüero, J. M., and M. S. Marks. 2011. "Motherhood and Female Labor Supply in the Developing World: Evidence From Infertility Shocks." *Journal of Human Resources* 46, no. 4: 800–826. https://doi.org/10.1353/jhr.2011.0002.

Alesina, A., and N. Fuchs-Schündeln. 2007. "Good-Bye Lenin (Or Not?): The Effect of Communism on People's Preferences." *American Economic Review* 97, no. 4: 1507–1528. https://doi.org/10.1257/aer.97.4.1507.

Angrist, J., and W. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence From Exogenous Variation in Family Size." *American Economic Review* 88, no. 3: 450–477. https://www.nber.org/papers/w5778.

Angrist, J. D., and G. W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity." *Journal of the American Statistical Association* 90, no. 430: 431–442. https://doi.org/10.2307/2291054.

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91, no. 434: 444–455. https://doi.org/10.1080/01621459.1996.10476902.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Ashcraft, A., I. Fernández-Val, and K. Lang. 2013. "The Consequences of Teenage Childbearing: Consistent Estimates When Abortion Makes Miscarriage Non-Random." *Economic Journal* 123, no. 571: 875–905. https://doi.org/10.1111/ecoj.12005.

Baker, M., J. Gruber, and K. Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116, no. 4: 709–745. https://doi.org/10.1086/591908.

Baranowska-Rataj, A., and A. Matysiak. 2016. "The Causal Effects of the Number of Children on Female Employment – Do European Institutional and Gender Conditions Matter?" *Journal of Labor Research* 37, no. 3: 343–367. https://doi.org/10.1007/s12122-016-9231-6.

Baranowska-Rataj, A., and A. Matysiak. 2022. "Family Size and Men's Labor Market Outcomes: Do Social Beliefs About Men's Roles in the Family Matter?" *Feminist Economics* 28, no. 2: 93–118. https://doi.org/10.1080/13545701.2021.2015076.

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, no. 1: 5–27. https://doi.org/10.1007/s00148-011-0370-z.

Becker, G. S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." Supplement, *Journal of Labor Economics* 3, no. S1, Part 2: S33–S58. https://doi.org/10.1086/298075.

Becker, G. S. 1991. *A Treatise on the Family*. enl. ed. Harvard.

Bhalotra, S., and D. Clarke. 2019. "Twin Birth and Maternal Condition." *Review of Economics and Statistics* 101, no. 5: 853–864. https://doi.org/10.1162/rest_a_00789.

Boeckmann, I., J. Misra, and M. J. Budig. 2015. "Cultural and Institutional Factors Shaping Mothers' Employment and Working Hours in Postindustrial Countries." *Social Forces* 93, no. 4: 1301–1333. https://doi.org/10.1093/sf/sou119.

Boelmann, B., Raute, A., and Schonberg, U. (2021). Wind of Change? Cultural Determinants of Maternal Labor Supply. CEPR Discussion Paper No. DP16149.

Braakmann, N., and J. Wildman. 2016. "Reconsidering the Effect of Family Size on Labour Supply: The Twin Problems of the Twin Birth Instrument." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179, no. 4: 1093–1115. https://doi.org/10.1111/rssa.12160.

Bronars, S. G., and J. Grogger. 1994. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *American Economic Review*: 1141–1156.

Bundesagentur Für Arbeit (BA). 2020. "Beschäftigungsquoten (Jahreszahlen und Zeitreihen)." *Bundesagentur für Arbeit, Statistik, Nürnberg*.

Campa, P., and M. Serafinelli. 2019. "Politico-Economic Regimes and Attitudes: Female Workers Under State Socialism." *Review of Economics and Statistics* 101, no. 2: 233–248. https://doi.org/10.1162/rest_a_00772.

Chhaochharia, V., Ghosh, S., Niessen-Ruenzi, A., and Schneider, C. (2021). Public Child Care Provision and Women's Career Choices. SSRN 2943427.

Collischon, M., Eberl, A., and Reichelt, M. (2020). Structural Legacies and the Motherhood Penalty: How past Societal Contexts Shape Mothers' Employment Outcomes in Reunified Germany. Working paper.

Cortés, P., G. Koşar, J. Pan, and B. Zafar. 2022. *Should Mothers Work? How Perceptions of the Social Norm Affect Individual Attitudes Toward Work in the U.S. Technical Report*. National Bureau of Economic Research.

D. I. R., Deutsches IVF-Register e.V. 2020. "Jahrbuch 2019: Deutsches IVF-Register." *Journal of Reproductive Medicine and Endocrinology* 17, no. 5. https://www.deutsches-ivf-register.de/perch/resources/dir-jahrbuch-2019-de.pdf.

Ellingsæter, A. L., R. H. Kitterød, and J. Lyngstad. 2017. "Universalising Childcare, Changing Mothers' Attitudes: Policy Feedback in Norway." *Journal of Social Policy* 46, no. 1: 149–173. https://doi.org/10.1017/s0047279416000349.

European Commission. 2024. *Directorate-General for Education, Youth, Sport and Culture, Monitor für die allgemeine und berufliche Bildung 2024 – Deutschland*. Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/c2ff23d6-ac70-11ef-acb1-01aa75ed71a1.

Eurostat. 2012. "Urban-Rural Europe - Labour Market." *Statistics Explained*. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/112338.pdf.

Farbmacher, H., R. Guber, and J. Vikström. 2018. "Increasing the Credibility of the Twin Birth Instrument." *Journal of Applied Econometrics* 33, no. 3: 457–472. https://doi.org/10.1002/jae.2616.

Federal Statistical Office 2024. *Frauen nach Zahl der Geborenen Kinder: Endergebnisse des Mikrozensus 2022*. Statistisches Bundesamt.

Federal Ministry of Finance. 2024. Datensammlung für Steuerpolitik. Bundesministerium der Finanzen, Referat für Öffentlichkeitsarbeit & Bürgerdialog. https://www.bundesfinanzministerium.de/Content/DE/Downloads/Broschueren_Bestellservice/datensammlung-zur-steuerpolitik-2024.pdf?__blob=publicationFile&v=5.

Federal Statistical Office. 2019a. *Kinderlosigkeit, Geburten und Familien - Ergebnisse des Mikrozensus 2018. Ausgabe 2019*. Statistisches Bundesamt.

Federal Statistical Office. 2019b. Lohn- und Einkommensteuerstatistik 2018, KDFV, Dataset of the Research Data Centres of the Federal

Statistical Office (Destatis) and the Statistical Offices of the Federal States. https://doi.org/10.21242/73111.2018.00.00.1.1.0.

Federal Statistical Office 2022a. *Statistiken der Kinder- und Jugendhilfe*. Statistisches Bundesamt.

Federal Statistical Office. 2022b. *Zusammengefasste Geburtenziffer nach Bundesländern (Kinder je Frau), 1990 – 2022*. Statistisches Bundesamt.

Federal Statistical Office. 2024a. "Durchschnittliches Alter der Mutter bei der Geburt: Deutschland, Jahre, Lebendgeburtenfolge." *Tabelle 12612-0014*. https://www-genesis.destatis.de/genesis/online.

Federal Statistical Office. 2024c. "Lebendgeborene: Deutschland, Jahre, Geschlecht." *Tabelle 12612-0001*. https://www-genesis.destatis.de/genesis/online.

Federal Statistical Office. 2024d. "Mehrlingsgeburten: Deutschland, Jahre, Art der Mehrlingsgeburt." *Tabelle 12612-0016*. https://www-genesis.destatis.de/genesis/online.

Fellman, J., and A. W. Eriksson. 2009. "On the History of Hellin's Law." *Twin Research and Human Genetics* 12, no. 2: 183–190. https://doi.org/10.1375/twin.12.2.183.

Fluchtmann, J. and V. Patrini. 2023. Joining Forces for Gender Equality: Women at Work in OECD Countries.

Fontaine, I. 2017. "The Causal Effect of Family Size on Mother's Labor Supply: Evidence From Reunion Island and Mainland France." In *Technical Report*: Working Paper.

Frenette, M. 2011. "How Does the Stork Delegate Work? Childbearing and the Gender Division of Paid and Unpaid Labour." *Journal of Population Economics* 24, no. 3: 895–910. https://doi.org/10.1007/s00148-010-0307-y.

Fuchs-Schündeln, N., and P. Masella. 2016. "Long-Lasting Effects of Socialist Education." *Review of Economics and Statistics* 98, no. 3: 428–441. https://doi.org/10.1162/rest_a_00583.

Grave, B. S., and C. M. Schmidt. 2012. "The Dynamics of Assortative Mating in Germany." *Ruhr Economic Paper*: 346.

Hakim, C. 2003. "A New Approach to Explaining Fertility Patterns: Preference Theory." *Population and Development Review* 29, no. 3: 349–374. https://doi.org/10.1111/j.1728-4457.2003.00349.x.

Hellwagner, T., D. Söhnlein, S. Wanger, and E. Weber. 2022. "Wie sich eine demografisch bedingte Schrumpfung des Arbeitsmarkts noch abwenden lässt." In *IAB-Forum, Nürnberg: IAB*. https://doi.org/10.48720/IAB.FOO.20221121.01.

Huber, K., and G. Rolvering. 2023. *Public Child Care and Mothers' Career Trajectories*. IZA Discussion Paper.

Jessen, J. 2022. "Culture, Children and Couple Gender Inequality." *European Economic Review* 150: 104310. https://doi.org/10.1016/j.euroecorev.2022.104310.

Karbownik, K., and M. Myck. 2016. "For Some Mothers More Than Others: How Children Matter for Labour Market Outcomes When Both Fertility and Female Employment Are Low." *Economics of Transition* 24, no. 4: 705–725. https://doi.org/10.1111/ecot.12104.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. 2019. "Child Penalties Across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109: 122–126. https://doi.org/10.1257/pandp.20191078.

Kleven, H., C. Landais, and J. E. Søgaard. 2019. "Children and Gender Inequality: Evidence From Denmark." *American Economic Journal: Applied Economics* 11, no. 4: 181–209. https://doi.org/10.1257/app.20180010.

Klinkhammer, N., B. Riedel, et al. 2018. *An Incomplete Revolution? Changes and Challenges Within German Early Childhood Education and Care Policy*, 49–70. SAGE handbook of early childhood policy.

Krause, P. 2019. "30 Jahre seit dem Mauerfall: Fortschritte und Defizite bei der Angleichung der Lebensverhältnisse in Ost-und Westdeutschland." *DIW Wochenbericht* 86, no. 45: 827–838. https://www.econstor.eu/handle/10419/206701.

Kremer, M. 2007. *How Welfare States Care: Culture, Gender and Parenting in Europe*. Amsterdam University Press.

Leibowitz, A., J. A. Klerman, and L. J. Waite. 1992. "Employment of New Mothers and Child Care Choice: Differences by Children's Age." *Journal of Human Resources* 27, no. 1: 112–133. https://doi.org/10.2307/145914.

Lippmann, Q., A. Georgieff, and C. Senik. 2020. "Undoing Gender With Institutions: Lessons From the German Division and Reunification." *Economic Journal* 130, no. 629: 1445–1470. https://doi.org/10.1093/ej/uez057.

Lundborg, P., E. Plug, and A. W. Rasmussen. 2017. "Can Women Have Children and a Career? IV Evidence From IVF Treatments." *American Economic Review* 107, no. 6: 1611–1637. https://doi.org/10.1257/aer.20141467.

Mätzke, M. 2010. "The Role of Old Ideas in the New German Family Policy Agenda." *German Policy Studies* 6, no. 3. https://d1wqtxts1xzle7.cloudfront.net/79348382/file-libre.pdf?1642868868=&response-content-disposition=inline%3B+filename%3DThe_Role_of_Old_Ideas_in_the_New_German.pdf&Expires=1760377857&Signature=HH1lfSgCWREIDfShLvcFOe5~hhsJPHpn31wPukWgZ6eyiN-TIMkSUKT1bSbLvxKPiASemOR6s7pbfi7Y~RJkzRtniIeu1~eNAHHenmBbPUlj~OOdGv4ilhSd2kos-y0lw8y8A7tiYYCQ71ZlYbU3sK1zRraTiwbEOOaixwjH~MmbNXX9t5ZpNxI17vM-TsckZLHP9bVomxyvvbB00hfmx9CZAlwKqIs1oYuamF7XtWL-pCQ5C4vgBL2LHtyA7guIENJKX0n7xnf9SE~1rBrIGpMW53uk4~baBJK0-j7g4N5TLeZ48bIcdgjVPgMx6-HV40temOq1EwjMdUP8U1JCuw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA.

Mincer, J. 1962. *Labor Force Participation of Married Women: A Study of Labor Supply*. Aspects of Labor Economics/NBER Books.

Moschion, J. 2013. "The Impact of Fertility on Mothers' Labour Supply in Australia: Evidence From Exogenous Variation in Family Size." *Economic record* 89, no. 286: 319–338. https://doi.org/10.1111/1475-4932.12042.

Olivetti, C., and B. Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics* 8, no. 1: 405–434. https://doi.org/10.1146/annurev-economics-080614-115329.

Pfau-Effinger, B. 2012. "Women's Employment in the Institutional and Cultural Context." *International Journal of Sociology & Social Policy* 32, no. 9/10: 530–543. https://doi.org/10.1108/01443331211257634.

Pfau-Effinger, B., and B. Geissler. 2002. "Cultural Change and Family Policies in East and West Germany." In *Analysing Families: Morality and Rationality in Policy and Practice*, edited by A. Carling, S. Duncan, and R. Edwards, 77–83.

Rønsen, M., and M. Sundström. 2002. "Family Policy and After-Birth Employment Among New Mothers–A Comparison of Finland, Norway and Sweden." *European Journal of Population/Revue europeenne de demographie* 18, no. 2: 121–152. https://doi.org/10.1023/a:1015532305179.

Rosenfeld, R. A., H. Trappe, and J. C. Gornick. 2004. "Gender and Work in Germany: Before and After Reunification." *Annual Review of Sociology* 30, no. 1: 103–124. https://doi.org/10.1146/annurev.soc.30.012703.110531.

Rosenzweig, M. R., and K. I. Wolpin. 1980. "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment." *Econometrica: Journal of the Econometric Society* 48, no. 1: 227–240. https://doi.org/10.2307/1912026.

Rosenzweig, M. R., and J. Zhang. 2009. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's 'One-Child' Policy." *Review of Economic Studies* 76, no. 3: 1149–1174. https://doi.org/10.1111/j.1467-937x.2009.00563.x.

Roth, A., and M. Slotwinski. 2020. "Gender Norms and Income Misreporting Within Households." In *ZEW-Centre for European Economic Research Discussion Paper. 20-001 (20-001)*.

Selin, H. 2014. "The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971." *International Tax and Public Finance* 21, no. 5: 894–922. https://doi.org/10.1007/s10797-013-9283-y.

Sigismund, M. 2018. *Regionalstatistische Raumtypologie (Regiostar) des BMVI für die Mobilitäts- und Verkehrsforschung*. BMVI. https://www.bmvi.de/SharedDocs/DE/Anlage/G/regiostar-arbeitspapier.pdf.

Silles, M. A. 2015. "The Impact of Children on Women's Labour Supply and Earnings in the UK: Evidence Using Twin Births." *Oxford Economic Papers* 68, no. 1: 197–216. https://doi.org/10.1093/oep/gpv055.

Spiess, C. K., and A. Dunkelberg. 2009. "The Impact of Child and Maternal Health Indicators on Female Labor Force Participation After Childbirth: Evidence for Germany." *Journal of Comparative Family Studies* 40, no. 1: 119–138. https://doi.org/10.3138/jcfs.40.1.119.

Trappe, H. 1996. "Work and Family in Women's Lives in the German Democratic Republic." *Work and Occupations* 23, no. 4: 354–377. https://doi.org/10.1177/0730888496023004003.

Trappe, H., and R. A. Rosenfeld. 2000. "How Do Children Matter? A Comparison of Gender Earnings Inequality for Young Adults in the Former East Germany and the Former West Germany." *Journal of Marriage and Family* 62, no. 2: 489–507. https://doi.org/10.1111/j.1741-3737.2000.00489.x.

Vere, J. P. 2008. "Dragon Children: Identifying the Causal Effect of the First Child on Female Labour Supply With the Chinese Lunar Calendar." *Oxford Bulletin of Economics & Statistics* 70, no. 3: 303–325. https://doi.org/10.1111/j.1468-0084.2008.00501.x.

Vere, J. P. 2011. "Fertility and Parents' Labour Supply: New Evidence From US Census Data: Winner of the OEP Prize for Best Paper on Women and Work." *Oxford Economic Papers* 63, no. 2: 211–231. https://doi.org/10.1093/oep/gpr003.

Witte, J. C., and G. G. Wagner. 1995. "Declining Fertility in East Germany After Unification: A Demographic Response to Socioeconomic Change." *Population and Development Review* 21, no. 2: 387–397. https://doi.org/10.2307/2137500.

Wu, X. 2022. "Fertility and Maternal Labor Supply: Evidence From the New two-Child Policies in Urban China." *Journal of Comparative Economics* 50, no. 2: 584–598. https://doi.org/10.1016/j.jce.2022.01.002.

Zhang, J. 2017. "A Dilemma of Fertility and Female Labor Supply: Identification Using Taiwanese Twins." *China Economic Review* 43: 47–63. https://doi.org/10.1016/j.chieco.2016.12.005.

Zoch, G., and P. S. Schober. 2018. "Public Child-Care Expansion and Changing Gender Ideologies of Parents in Germany." *Journal of Marriage and Family* 80, no. 4: 1020–1039. https://doi.org/10.1111/jomf.12486.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.
**Supporting Information S1:** ecot70014-sup-0001-Appendix.pdf.