Child care expansion and mothers' labor supply: is there a causal link?

Kai-Uwe Müller, Denise Sengül, Katharina Wrohlich*

This version: February 11, 2015

Preliminary and incomplete!

Please do not quote without permission of the authors!

This paper assesses whether there is a causal link between the provision of publicly subsidized childcare and the labor supply of mothers. We contribute to the related quasi-experimental literature by focusing on mothers with children aged 1 to 3. The effects of full-time and part-time care are disentangled. We exploit spatial and temporal variation in the expansion of publicly subsidized childcare triggered by comprehensive policy reforms. The utilization of various data sets and a systematic comparison of estimation frameworks sheds light on this relationship under different identifying assumptions. The crucial point is whether identification is restricted to quasi-experimental variation within regions. We confirm previous findings by showing the sensitivity of results to the choice of the research design, in particular the source of variation. Relying on credible exogenous variation we do not find a significant impact of childcare expansion on mothers' extensive labor supply margin. We find, however, a significant effect at the intensive margin. Our results cast doubt on previous empirical findings in terms of identification and effect size.

Keywords: childcare provision; mother's labor supply; instrumental variables, generalized difference-in-difference

JEL classification: J22; J13; H43;

* Corresponding author: German Institute for Economic Research Berlin (DIW Berlin), Mohrenstr. 58, D-10117 Berlin, Tel.: +49 30 89789 164, Fax: +49 30 89789 114, E-mail: kwrohlich@diw.de

Contents

1 Introduction

In light of low fertility and female employment rates the provision of public childcare has been on the agenda of many countries (Immervoll and Barber, 2006). The expansion of public childcare is supposed to increase fertility, mothers' labor market attachment, and promote childrens' development in early life. A positive effect of public childcare on maternal employment can be rationalized within an economic model of the family (Blau, 2003). The implied decrease in costs of childcare changes the relative utilities of consumption and leisure. Yet, income effects, preferences for the quality of care, or the availability of alternative modes of private care loosen this relationship. Subsidized childcare may crowd-out other forms of (private or informal) care conditional on maternal employment.

Given the theoretical ambiguity, an empirical literature on the effects of subsidized childcare has emerged. Evidence from early reduced-form studies and structural models unequivocally points to a significant impact of costs and availability of childcare on mothers' labor supply (Anderson and Levine, 2000; Blau and Currie, 2006). A different type of approach exploits quasi-experimental variation induced by policy reforms. Results from those studies are mixed. Although a majority of findings confirms the significant effect of childcare,[1] several analyses report negligible or insignificant estimates.[2] Moreover, effects are often found to be heterogenous in terms of family status, additional children in the household, childrens' age, and mothers' qualification (see, e.g., Cascio, 2009 or Fitzpatrick, 2012).

Our study makes use of two comprehensive policy reforms in Germany from 2005 and 2008. Their implemention created a suitable quasi-experimental setting. Starting from very low levels, childcare coverage for children aged 0 to 3 in West Germany more than doubled from about 8% to more than 20%. We are thus able to work with a substantial increase in the provision of childcare and not only marginal changes in childcare costs. At the same time the level of public care for children aged 4 to 6 remained largely constant (see Fig. A1 in the Appendix). Peculiarities of

[1]See Gelbach (2002), Blau and Tekin (2007), Berlinski et al. (2011), and Felfe et al. (2014) for examples using an instrumental variables (IV) approach and Berger and Black (1992), Berlinski and Galiani (2007), Baker et al. (2008), Lefebvre and Merrigan (2008), Lefebvre et al. (2009), Simonsen (2010), Havnes and Mogstad (2011), Nollenberger and Rodríguez-Planas (2011), and Schlosser (2011) for estimates from a difference-in-difference (DD) framework.

[2]See Fitzpatrick (2010) for an examplaric IV study and Lundin et al. (2008) and Havnes and Mogstad (2011) for the DD approach.

the adiministrative process and frictions in the market for childcare led to regional variation over time. The differences are generated across municipalities, albeit only observed at the county (*Kreis*) level in the data available to us. We argue that – conditional on covariates – (part of) this variation can be considered exogenous and used for identification. In this paper we assess whether there is a causal link between the provision of publicly subsidized childcare and the labor supply of mothers with children aged 1 to 3 in West Germany.

Quasi-experimental approaches sidestep some identification issues arising in structural or reduced-form estimations: decisions on a child care arrangement and mother's labor supply are, for instance, not made independently. Costs (sometimes the availability) of informal care are not observed. Instead, variation in the costs or provision of childcare is used that is generated by processes exogenous to mothers' employment and childcare choices. Our study contributes to the branch of this literature which is based on regional variation in the supply of childcare (see, e.g., Havnes and Mogstad, 2011). The main problem here is whether (or which portion of) it is truly exogenous. Depending on the assumptions made different estimation strategies can be applied.

Several papers on Germany have exploited the cross-sectional part of this variation. It serves as exclusion restriction for the determination of childcare costs in structural models (Wrohlich, 2011; Haan and Wrohlich, 2011; Müller and Wrohlich, 2014), as excluded instrument in instrumental variable (IV) frameworks (Bauernschuster and Schlotter, 2013; Felfe and Lalive, 2013) or is used directly in reduced-form employment equations (Kreyenfeld and Hank, 2000; Büchel and Spie"s, 2002; Schober and Spie"s, 2014). Cross-sectional differences in the supply of childcare may be endogenous, however: Parents with given preferences could demand a certain amount of childcare. Childcare providers (parents) could select into a municipality with a specific demand for (supply of) childcare. Municipalities could use the provision of childcare to attract high quality labor (Felfe et al., 2014). Relying only on within-region variation may be less restrictive when the expansion of childcare as a result of policy reforms is subject to implementation frictions.

In this paper we systematically compare different empirical approaches that rely on alternative identifying assumptions. We use rich household data from the German Socio-Economic Panel (SOEP) as well as several waves of the German Microcensus

(MC) to avoid size limitations of the SOEP. Information on subsidized childcare and control variables at the county level are provided by the German Statistical Office and the German Youth Institute (Deutsches Jugendinstitut e.V., DJI). The SOEP contains the mother's labor market status and the choice of a care arrangement. We are thus able to estimate local average treatment effects (LATE) within an IV framework. Contrasting specifications without and with county fixed effects reveals that the identifying variation is not sufficient in the latter.

We therefore turn to the large MC providing more variation within counties. It does not contain information on the individual choice of a child care arrangement, though. Therefore we estimate intention-to-treat effects (ITT) of the provision of public childcare on maternal labor supply. The identifying assumptions in terms of county fixed effects, i.e. reliance on within-county vs. overall variation, turn out to be crucial for the estimated effects.

While most previous work is focused on labor force participation, we distinguish various indicators of maternal labor supply, including the intensive margin. We are furthermore able to discriminate between the provision of full-time and part-time care. Our preferred specifications include control variables at the individual and regional level that might affect mothers' labor supply and childcare decision. The estimates are further conditioned on regional determinants potentially driving the local supply of childcare.

We find significantly positive effects of subsidized childcare on maternal labor supply based on the overall variation in childcare provision. This holds for all measures of labor supply at the extensive and intensive margin as well as IV and reduced-form frameworks. This picture, however, changes dramatically when identification is restricted to quasi-experimental within-county variation. Childcare coverage indicators turn out to be weak instruments in estimations with the SOEP; the childcare effect becomes insignificant. The difference-in-difference framework which exploits the same type of variation also yields insignificant estimates for all outcome variables. Using the larger MC data with more within-county variation and employing a more efficient regression approach with county fixed effects we find marginally significant effects at the intensive margin.

In line with previous assessments (Blau, 2003; Blau and Currie, 2006) we conclude that estimates based on quasi-experimental variation are very sensitive with respect to identifying assumptions and related model specifications. This has far-reaching consequences for the substantive implications of the findings. It casts doubt on previous estimates of significant participation effects – most of them did neither relies exclusively on quasi-experimental variation, nor provide robustness checks with county fixed effects.

Taking the findings based on the arguably more credible exogenous quasi-experimental variation at face value there is no identifiable extensive margin effect of the expansion of child care for children aged 1 to 3 on their mothers' labor supply. Already employed mothers may increase their working hours, though. Comparable studies for Sweden and Norway found similar patterns. It furthermore confirms small extensive margin elasticities of structural approaches. Showing that there might be a systematic difference between the extensive and intensive margin adds to the international literature. As does the focus on mothers with young children and the flexible specification in terms of full- and part-time childcare provision.

The rest of the paper is organized as follows. Section 2 provides some theoretical considerations about the relationship between the provision of publicly subsidized childcare and maternal employment, reviews the related empirical literature, and discusses developments on the German child care market. Section 3 gives an overview on the different datasets used. The empirical strategies are discussed in section 4 including identification issues, variable and sample definitions. Empirical results are presented in section 5 including a number of robustness checks. Section 6 summarizes the main findings and concludes.

2 Provision of child care and maternal labor supply

2.1 Theoretical considerations

We interpret the causal impact of public childcare on the employment decision of mothers within an economic framework of the family (Becker, 1981; Blau, 2001). Mothers in single or couple households (co-)decide simultaneously on their labor supply and a care arrangement for their children by maximizing a utility function in the arguments of consumption, leisure, and quality of care subect to bugdet and time constraints. In terms of childcare modes a complete choice set includes maternal care, (unpaid) informal care by relatives or friends, private formal care (by nannies or for-profit providers), and formal care in publicly subsidized/financed care centers. Households might be constrained with respect to informal as well as public childcare whereas private care can always be obtained at market prices.

The provision (subsidization) of public childcare affects the budget constraint by reducing costs for this form of care. The number of alternatives in the choice set of households constrained in certain care dimensions increases. As a consequence absolute and relative prices for different modes of care are altered wich affects utitlity in the corresponding alternatives. The substitution effect leads ceteris paribus to a higher utilization of public childcare and increased maternal labor supply. This is the main channel for the supposed positive relationship between subsidized childcare and mothers' employment. As Blau (2003) points out the associated income effect goes in the opposite direction. It depends on her preferences which effect dominates and whether a given woman will raise or reduce labor supply. This is one of the sources of ambiguity or heterogeneity in empirical estimates (Gelbach, 2002; Cascio, 2009).[3]

Including (unpaid) informal care in the analysis provides another margin of adjustment. Blau (2003) notes that the provision of subsidized childcare changes relative costs of formal and informal care conditional on employment. Likewise households might substitute between different modes of care with the labor supply of mothers remaining constant. This is one of the mechanisms cited regularly in the empirical literature for explaining small or insignificant estimates for the effect of childcare on

[3]This ambiguity is particularly relevant for the intensive labor supply margin and larger at in case of non-linear subsidies (Blau, 2003).

maternal labor supply (Havnes and Mogstad, 2011). Felfe and Lalive (2013) provide some evidence for this channel.

Childcare quality and parental preferences for different modes of care related to children's well-being also influence behavioral reactions. Public childcare will only be utilized when its perceived quality is sufficiently high compared to the alternatives of maternal (or informal) care. It is conceivable – particularly for very young children – that monetary incentives do not outweigh quality considerations. This constitutes another mechanism that impedes an effect of public childcare on maternal employment.[4]

Fitzpatrick (2010) refers to another point not related to childcare. Rationed labor demand or working hours constraints can also be reasons why improved labor supply incentives do not translate into higher employment. Moreover, the strength of the relationship between childcare and mothers' employment will depend on the context. Preferences and care cultures vary considerably across, sometimes even within countries. Germany is one example with traditionally stark contrasts between the West and the East in utilization of early care and mothers' employment rates.

The different reasons for a loose link between childcare and employment can be qualified in terms of effect heterogeneity related to incentives or preferences. The aforementioned factors might be only binding for different sub-groups. Monetary incentives often vary between single and married or cohabiting mothers. Mothers' decisions related to a given child will be influenced by (younger) siblings. The latter point is related to preference heterogeneity with respect to the age of a child. Mothers are in general less willing to send their children to public childcare at very young ages (Pungello and Kurtz-Costes, 1999). Womens' level of qualification relates to heterogeneity in incentives and preferences. Empirical evidence supports these different aspects (Anderson and Levine, 2000; Blau and Currie, 2006).

To sum up, there are several reasons why the impact of publicly subsidized childcare on maternal employment might be ambiguous. Whether or not there is a significant effect depends on the empirical application. In addition to economic incentives and preferences, to the market for childcare and 'care culture', it depends on the population of mothers and children considered.

[4] Another (rather extreme) case is that mothers only work to finance high quality childcare. When the price for this type of care is reduced by public subsidies, these mothers might reduce their employment.

2.2 Empirical literature

Empirical studies of the relationship between public(ly subsidized) childcare and mothers' labor supply exploit different sources of variation. Childcare costs may vary at the household level, across regions, or as a result of childcare policy (reform). Using individual or regional variation of childcare costs reduced-form and structural studies estimate employment elasticities for mothers. Structural models are also used to simulate the outcomes of existing or proposed policies. Evaluation studies exploit quasi-experiments that generate exogenous variation in the costs or provision of public childcare.[5]

Early empirical studies are based on *reduced-form* employment *regressions* with the utilization of public childcare as main explanatory variable, The main methodological challenges are the endogeneity of childcare in this estimation as well as selection problems related to employment and public childcare. The literature is dominated by studies on the U.S. (Blau and Robins, 1991; Connelly, 1992; Ribar, 1992; Kimmel, 1995; Powell, 1997; Kimmel, 1998; Anderson and Levine, 2000; Han and Waldfogel, 2001; Baum II, 2002; Meyers et al., 2002). Virtually all studies find a significantly relationship between costs of childcare and maternal employment with a considerable range of estimates ranging between elasticities close to zero and below -3. Blau and Currie (2006) argue that this can be explained by methodological dicrepancies (specification, exclusion restrictions, controls) and much less by different data sources and samples.

Structural approaches directly model the decisions on a child care arrangement and maternal labor supply. Simultaneity of the care and employment choices, selectivity issues, different modes of care, and rationing on the childcare market are addressed within this framework. The first and most of the studies again refer to the U.S. (Heckman, 1974; Blau and Robins, 1988; Michalopoulos et al., 1992; Ribar, 1995; Averett et al., 1997; Blau and Hagy, 1998; Michalopoulos and Robins, 2002; Connelly and Kimmel, 2003; Tekin, 2007). More recently structural evidence is also available for Sweden (Gustafsson and Stafford, 1992), the UK (Duncan and Giles, 1996; Blundell et al., 2000a,b; Duncan et al., 2001; Parera-Nicolau and Mumford,

[5]See Blau (2003) and Blau and Currie (2006) for an overview of comparable methodological approaches on other outcomes like utilization of childcare or childrens' development. Outcomes for women include fertility (see Haan and Wrohlich (2011) for an exemplary study on Germany) or welfare receipt (see, e.g., Connelly and Kimmel, 2003).

2005), Canada (Michalopoulos and Robins, 2000; Powell, 2002; Michalopoulos and Robins, 2002), Italy (Del Boca, 2002; Del Boca and Vuri, 2007; Del Boca and Sauer, 2009), and France (Choné et al., 2003). Although smaller in magnitude compared to reduced-form work, the vast majority of those studies also finds significant employment elasticities. Main challanges in this framework include the endogenous access to different types of care with respect to mothers' employment and the unobserved costs or availability of informal care.

Quasi-experimental settings generated by policy reforms have been used increasingly in more recent years to circumvent some of the identification issues.[6] When individual information on the utilization of subsidized childcare is available, the exogenous variation is used to instrument the childcare choice within an IV framework. This variation is generated either by birth thresholds for the enrollment into childcare or preschool programs (Gelbach, 2002; Fitzpatrick, 2010; Goux and Maurin, 2010; Berlinski et al., 2011; Fitzpatrick, 2012). The staggered introduction or expansion of subsidies (provision of public care) across regions serves as alternative instrument (Blau and Tekin, 2007; Felfe et al., 2014). We use the latter identification strategy in this paper (sub-section 4.1).

Without household information on the choice of childcare, quasi-experimental variation is used in a DD or panel framework to identify intention-to-treat effects. Some studies exploit exogenous variation within a single state (Berger and Black, 1992) or the difference between a single treated region vs. the rest of the country (Baker et al., 2008; Lefebvre and Merrigan, 2008; Lefebvre et al., 2009). In the most common setting, childcare or preschool policies induce exogenous variation between and within several regions of a country. The effect can then be estimated in a (generalized) DD design with a region-time-specific treatment indicator, region and time fixed effects and control variables at the individual and regional level. Evidence is available for a number of different institutional contexts; see Cascio (2009) for the U.S., Berlinski and Galiani (2007) for Argentina, Lundin et al. (2008) for Sweden, Simonsen (2010) for Denmark, Havnes and Mogstad (2011) for Norway, and Schlosser (2011) for Israel. Most of the empirical work of this paper is conducted within this framework (sub-sections 4.2, 4.3).

[6]Yet another branch of the literature is based on social experiments and demonstration projects; see Gennetian et al. (2001), Blau (2003), Blau and Currie (2006), and Blau and Tekin (2007) for further references.

The quasi-experimental literature has put earlier findings somewhat into perspective. Results are often heterogenous, in some instances statistically or economically insignificant (Fitzpatrick, 2010; Lundin et al., 2008; Havnes and Mogstad, 2011). Employment effects turn out to be more often significant for single than married or cohabiting mothers (Goux and Maurin, 2010; Cascio, 2009). The impact on maternal employment is also repeatedly found to be absent with younger siblings in the family (Gelbach, 2002; Berlinski et al., 2011; Cascio, 2009). Related to that effects tend to be higher for older children (Goux and Maurin, 2010).

The *evidence for Germany* is based on all three approaches. In two early studies cross-sectional variation in childcare coverage is included in reduced-form employment equations. Using the SOEP for 1996 Kreyenfeld and Hank (2000) do not get a significant effect for mothers with children below the age of 12 in West Germany. Büchel and Spie"s (2002) use SOEP data for 1998, restrict the sample to preschool children and include the share of full-time slots as additional regressor. They find more significant effects for part-time employment and larger estimates for older children. Schober and Spie"s (2014) use data from the SOEP and 'Families in Germany' ('Familien in Deutschland', FiD) for 2010/11. In addition to quantitative childcare indicators they include quality measures at the county level (available for one year) in an identical research design. They get insignificant quantity effects and partially significant coefficients for quality. Signs of those coefficients partially change between West and East Germany and are not always consistent with theoretical expectations.[7]

There are several structural models estimated on German data. Wrohlich (2011) develops a framework for the joint decisions on childcare and mothers' labor supply. The most important feature of this model estimated with SOEP data from 2001 to 2003 is that it deals with rationing in the childcare market by exploiting regional variation on childcare coverage. Wrohlich finds relatively low childcare cost elasticities of labor supply (-0.13). Müller and Wrohlich (2014) extend this framework by considering childcare choices of all children in the household below the age of 13 based on data from SOEP and FiD for 2010. They use individual information on access restrictions to childcare from FiD and include regional dummies in the

[7]We consider the specification of aggregate 'quality indicators' problematic. The effect of care quality is not identified at this level of aggregation by including these measures as separate regressors. This applies also to the share of full-time slots in Büchel and Spie"s (2002).

rationing equation. Their analysis yields even smaller overall childcare cost elasticities (-0.08). These findings are plausible given the significant rationing in the West German market for childcare (Wrohlich, 2008). Both studies simulate significantly positive labor supply effects of an expansion of public childcare. This finding is also supported by two dynamic structural models based on SOEP data. Haan and Wrohlich (2011) estimate a model that jointly determines a women's labor supply and fertility decision and includes feedback effects. Bick (2011) calibrates a life-cycle framework with a detailed depiction of childcare decisions. Both studies simulate substantial employment effects as a consequence of an extended provision of care.

Most closely related to this paper are studies that exploit quasi-experimental variation at the regional (county) level.[8] Felfe and Lalive (2013) use regional variation in the provision of publicly subsidized care for children aged 0 to 3 as an instrument for the endogenous choice of childcare in the first stage of a marginal treatment effects framework. In their estimation based on SOEP data for the years 2002 to 2008 they include fixed effects at the state, not at the county level. Although being primarily interested in child outcomes, they also look at labor supply and find substantial positive effects at the extensive margin. Bauernschuster and Schlotter (2013) analyze the effect of publicly subsidized childcare for 3 to 4 year olds on their mothers' employment based on SOEP data from 1991-2005. They exploit regionally varying cut-off rules for the access to a kindergarten place during the implementation of a reform-induced expansion of care. Although they do not observe the actual distribution of rules, they take the eligibility criterion as an instrument in the first stage of an IV model. Similar to Felfe and Lalive (2013) they control for state, but not for county fixed effects. According to their results a 10 percentage point increase in public childcare increases maternal employment 3.5 percentage pointes.

The work by Bauernschuster et al. (2013) is the only study for Germany that closely resembles the DD/panel design of our study (sub-sections 4.2, 4.3). They are interested on the impact the expansion of childcare for children in the age of 0 to 3 years had on fertility. A DD specification similar to Havnes and Mogstad (2011) is used where regions with an above-median increase in childcare coverage during the period of analysis are considered as treated. Alternatively a standard fixed effects

[8]See Coneus et al. (2008) and Bauernschuster and Schlotter (2013) for alternative instruments and Bauernschuster and Schlotter (2013) for an alternative control group.

linear panel model is estimated. They use administrative data aggregated at the county level and find a significantly positive effect – a 10 percentage point increase in care lead to 1.4 more births per 1000 women (3.2%).

How does our study line up with the literature? It is based on commonly used quasi-experimental variation at the regional level. In terms of treatment intensity the policy reform is midtable, comparable to Havnes and Mogstad (2011), above Lundin et al. (2008), and below Nollenberger and Rodríguez-Planas (2011). Although being standard in the international literature, accounting for regional fixed effects in a DD or fixed effects panel framework has not been done so far in empirical studies for Germany. We analyze the effects in the West-German context with low pre-reform levels of fertility, public childcare coverage and maternal employment. This might mostly resemble Nollenberger and Rodríguez-Planas (2011) and Havnes and Mogstad (2011). Moreover, we focus on children aged 1 to 3, which has not been so much in the focus of international studies that mostly look at children above the age of 3.

2.3 The child care market and policy reforms

Germany has traditionally been characterized by low fertility and labor force participation; the latter holds in particular for mothers with young children (Fig. A2 in the Appendix). Besides incentives of the tax and transfer system, social norms or attitudes towards motherhood and women's employment, the low supply of formal childcare is often quoted as an important cause. Peculiar for Germany is a stark regional contrast: Women and mothers in the East have been much better integrated into the labor market than their peers in the West. Due to the division of the country family policies have diverged historically. Childcare coverage has therefore been much lower in West Germany, in particular for children under the age of three (Fig. A1 in the Appendix).

Except for the legal claim to a kindergarten place in the child and adoloscent support law (*"Kinder- und Jugendhilfegesetz"*) of 1996, policy reforms have only been initiated since the middle of the 2000s (Spie"s, 2011). The day care expansion law (*Tagesbetreuungsausbaugesetz*, TaG) adopted in 2005 explicitly addresses the demand-oriented expansion of care for children under the age of three. The law particularly aims at enhancing the quality of care by childminders (*Tagespflege*). It formulates explicit quality standards to render these equivalent to alternative childcare facilities. In December 2008 the law on support for children (*Kinderförderungsgesetz*, KiFöG) was adopted that commits states to a gradual expansion of childcare supply for children under the age of three. A binding deadline was defined when local supply has to meet demand for childcare. As of 1 August 2013 each child under the age of three is legally entitled to a subsidized child care slot.[9]

West Germany has experienced a large expansion of publicly subsidized early child care. The increase started from an average level of about 2% in 2002 reaching roughly 22% in 2011 (Fig. A1 in the Appendix). East Germany, in comparison, has been characterized by a high supply of formal child care since German re-unification. Coverage for children aged 3 to 6 was already very high at the end of the 1990s.[10] The German care market for children under the age of three has been characterized by excess demand (Wrohlich, 2008). Although its degree has declined, rationing still persists as demand increased parallel to supply. According to newly available FiD data it amounts to 16% in West and 14% in East Germany (Müller and Wrohlich, 2014). We thus assume a full take-up of newly created childcare slots for children under the age of three for the subsequent empirical analysis.

Publicl childcare in Germany is provided by communities. Private providers include religious non-profit, non-religious non-profit, or commercial institutions. Public together with non-religious and religious non-profit providers cover almost the entire market; in 2009 roughly 2% of slots were owned by commercial providers (Mühler, 2010). Market composition varies across regions (Mühler, 2010; Hüsken, 2010, 2011). For children under the age of three, religious and commercial centers provide the majority of slots, especially for full-time care. The government also subsidizes certified child minders that take care of children outside of their homes. This comes at considerably higher cost. The share of public funding in the German childcare system is quite low compared to other European countries. Parents pay income-dependent fees with rules varying across municipalities.

[9] Children under the age of one only have this legal right, if their parents are working, currently searching for work or in education.

[10] The reduction of childcare provision for this age group in Fig. A1 is due to a change in data collection. Before 2006 availability was measured in slots per 100 children (by age group). It is defined as the percentage of children in child care since then. This did not affect figures for under three year olds because of rationing.

The expansion of subsidized child care following the aforementioned reforms is financed in part by the federal government and partly by the states. The amount and composition of funds vary by state-specific contracts. The general objective, strategy and funding are determined at the federal and state level. Youth welfare offices (*Jugendämter*) and/or municipal governments do most of the operational planning; arrangements vary between states (Hüsken, 2010, 2011).[11] Local authorities estimate the number of additional slots (and amount of daily care) needed and develop an appropriate expansion strategy. They refer to residents' registration statistics (*Einwohnermeldestatistiken*), use childcare facilities' waiting lists and information on past years. Almost half of the youth offices declared to interview parents with respect to their care preferences (BMFSFJ, 2011, 2012, 2013). This process is not codified by federal or state laws and subject to projection error. The vast majority of municipalities report financial, personell and spatial shortages when trying to meet local demand (BMFSFJ, 2011, 2012, 2013). There are thus substantial frictions in the implementation of those reforms. We exploit this type of variation arising from the reform-induced expansion of publicly subsidized childcare for identification (sub-section 4.4).

Two policy reforms which potentially influenced maternal labor supply coincided with the expansion of childcare during this period. The reform of unemployment insurance and assistance in 2005 by the so-called 'Hartz laws' comprised a large decrease in umemployment benefits. The implied higher labor supply incentives affected everyone receiving unemployment benefits (Jacobi and Kluve, 2007). The parental allowance reform (*Elterngeld*) reform from 2007 reduced the maximum duration of maternal (parental) benefits to 12 (14) months while relating the benefits to previous labor earnings. It only affected employment incentives of mothers in the first and second year after childbirth. The new parental leave scheme provides negative work incentives for women in the first year after childbirth and positive work incentives afterwards compared to the pre-reform period (Geyer et al., 2014). Both reforms applied at the federal level and were – at least not not by construction – systematically related to the expansion of subsidized childcare (sub-section 4.4).

[11] See Table **??** [ADD!] for an overview of the different federal laws and how they oblige whom.

3 Data

Since information on subsidized childcare is only available at the level of German counties, we resort to these regional data provided by the German Statistical Office or the German Youth Institute (Deutsches Jugendinstitut e.V., DJI). Additional indicators driving the local demand for childcare and affecting maternal labor supply are also available at this level of agregation (sub-section 3.1). The German Socio-Economic Panel (SOEP) provides detailed measures on the utilization of subsidized childcare and the labor market state of mothers at the individual and household levels (sub-section 3.2). The significantly larger German Microcensus (MC) that contains detailed information on the family and employment status of women (sub-section 3.3).

3.1 Regional data

Information on the supply of childcare are only available at the level of German counties (*Kreise*). For the years 1994, 1998, and 2002 these data were gathered by German Youth Institute based on material by the Statistical Offices of federal states. Provision of childcare was measured as slots available per 100 children in the respective age groups without information on the actual take-up of these places. From 2006-2012 data has been provided by the German Statistical Office. The indicator is defined as the percentage of children using subsidized formal child care in this period. For the age group under the age of three the difference between both concepts is not too severe because there has been substantial rationing in the market and take-up was therefore high.

From 2007 onwards we also can distinguish full-time and total covarage of subsidized childcare. We consider 2007 to 2011 as our baseline period of analysis. The MC wave for 2012 is not yet available to us. Robustness checks include data from earlier years. We have not exploited data for East Germany. There exists no consistent county panel due to reorganizations of local governments. We plan to include East Germany in the next version version of this paper on he basis of a harmonized county panel.

In addition, we use several control variables that are collected and edited jointly by the German Statistical Office with the Federal Institute for Research on Building, Urban Affairs and Spatial Development within the Federal Office for Building and Regional Planning. The dataset "Indicators and Maps on the Spatial Development" ("Indikatoren und Karten zur Raumentwicklung", INKAR, see Helmcke, 2008) allows longitudinal comparisons at different regional levels for Germany. The information used here is aggregated at the county level. Childcare data and INKAR are merged with SOEP and MC using hose county identifiers.

3.2 German Socio-Economic Panel

We first estimate our models based on a subsample of the *Socio-Economic Panel (SOEP)*. The SOEP is a representative longitudinal household study that started in 1984 and contains information on roughly 20,000 individuals living in 12,000 households in the year 2010 (Wagner et al., 2007). Mothers are diectly linked to their partner and children. It contains detailed information on the choice of a childcare arrangement for different children within a household (Wrohlich, 2011; Müller et al., 2014). Besides the utilization of publicly subsidized childcare we know whether family members or friends provide informal care (sub-section 4.5). Regional information on the provision of subsidized childcare as well as a range of regional control variables can be matched on the basis of county identifiers.

The SOEP also provides detailed information on the labor supply of mothers. In addition to the labor market status for all individuals, we know the type of employment relationship of those participating in the labor market, their (contractual and actual) weekly working hours as well as their labor earnings. Moreover, a wide range of individual and household characteristics is available that shed light on the economic and socio-demographic background of the mothers in our sample(s) (see sub-section 4.5 for details).

The main concern with this data set is that the number of observations might be too small for the identification of a causal effect. This problem has been documented in comparable studies (see, e.g. Gathmann and Sass, 2012; Felfe and Lalive, 2013). In our application it turns out to be particularly problematic in specifications with regional fixed effects where within-county variation is insufficient. Therefore we resort to a second, substantially larger data set, the German Microcensus.

3.3 German Microcensus

Given the limitations of the SOEP we utlize the German Microcensus (MC) which is the largest household survey of all European countries (Lengerer et al., 2005; Lotze and Breiholz, 2002a,a). Being a representative one percent sample of the German population it provides considerably more within-county variation at the individual level. The MC has a particular focus on demographic and labor market related topics. In this regard it closely resembles information which is available in the SOEP and can be observed and compared over all waves of the MC. It is thus a natural complement to the SOEP for our application. Comparable studies for Germany also use the MC for these reasons (Bauernschuster and Schlotter, 2013). The MC is collected annually and has a panel dimension for a sub-population over a limited period of time that is not exploited here. Since households are sampled as a whole, mothers are linked to their children.

Contrary to the SOEP, the MC does not contain information on the individual choice of a childcare arrangement. Yet, it provides comparable indicators for the extensive and intensive labor supply margin of individuals as well as the current contract status of those in employment. In addition, most of the individual-level control variables are available. In 2005 the survey design changed from the consideration of a fixed reference week per year to the collection of information during a year (Afentakis and Bihler, 2010). This affected particularly variables related to the employment status. Our baseline estimates therefore refer to the period after this conceptual break.

4 Empirical methodology

Depending on the data used and the variation available, different approaches to estimate the relationship between subsidized childcare and maternal labor supply become available. Should the mother's labor market status and her choice of a care arrangement be observed, instrumental variables estimation is the natural choice (sub-section 4.1). The endogenous choice of public childcare in the structural equation is instrumented by the exogenous quasi-experimental variation in childcare coverage at the county level. Given the validity of the assumptions we identify the local average treatment effect (LATE). Since the identifying variation is limited, particularly when regional fixed effects are introduced, the instrument turns out to be weak in most specifications. A larger data set that provides more variation does, unfortunately, not contain information on the individual choice of a child care arrangement. Under these circumstances we use the quasi-experimental regional variation in subsidized childcare to estimate an intention-to-treat effect (ITT) on maternal employment. Difference-in-difference and fixed effects estimatation is used to rule out time-invariant unobserved confounders affecting childcare and mothers' employment (sub-sections 4.2, 4.3). Sub-section 4.5 discusses the definition of variables and estimation samples.

4.1 Instrumental variables

When we observe the employment outcome y_{ijt} of mother i in region j at time t and her binary choice choice of childcare ca_{ijt}, a dummy that equals one for public care and zero otherwise, we can write y_{ijt} as a linear function of ca_{ijt} and other covariates:

$$y_{ijt} = \alpha + \delta ca_{ijt} + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \gamma_t + e_{ijt} \tag{1}$$

Assume X_{ijt} and X_{jt} are exogenous variables at the individual and regional level, γ_t are time fixed effects, and e_{ijt} is an error term. The variable ca_{ijt} in this equation is endogenous for at least two reasons. First, a mother's decisions about her labor supply and the form of childcare are taken simultaneously and might be influenced by the same factors. Second, mothers that send their children to public childcare may be different from others in terms of unobserved characteristics that might likewise affect their employment decision. In order to get a causal interpretation, ca_{ijt} is instrumented with quasi-experimental variation in the regional childcare coverage cc_{jt}:

$$ca_{ijt} = \eta + \kappa cc_{jt} + X'_{ijt}\varphi_1 + X'_{jt}\varphi_2 + \gamma_t + \upsilon_{ijt} \tag{2a}$$

$$ca_{ijt} = \eta + \kappa cc_{jt} + X'_{ijt}\varphi_1 + X'_{jt}\varphi_2 + \gamma_t + \mu_j + \upsilon_{ijt} \tag{2b}$$

Other studies have used a similar instruments (Bauernschuster and Schlotter, 2013; Felfe and Lalive, 2013; Blau and Tekin, 2007).[12] Two specifications are distinguished: (2a) exploits all variation in cc_{jt} conditional on observed covariates X and fixed time effects γ_t capturing overall trends. (2b), in addition, includes regional fixed effects μ_j at the county level. This specification has not been estimated in comparable IV studies.[13] Only within-county variation in public childcare coverage over time is utilized as instrument. The two corresponding specifications for the structural equation are as follows:

$$y_{ijt} = \alpha + \delta \hat{ca}_{ijt} + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \gamma_t + \varepsilon_{ijt} \tag{3a}$$

$$y_{ijt} = \alpha + \delta \hat{ca}_{ijt} + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \gamma_t + \mu_j + \varepsilon_{ijt} \tag{3b}$$

We estimate both stages by OLS facilitating the inclusion of individual fixed effects. Angrist (2001) discusses the adequacy of the linear probability model for binary outcomes which applies in our case to the first stage and binary outcomes of the second stage (sub-section 4.5 below).[14]

Instrument validity implies that the effect of cc_{jt} runs only through the choice of an individual care arrangement ca_{ijt}. Regional childcare coverage should have no other, indirect effects on mothers' employment. An obvious mechanism that violates this assumption is that the supply of childcare is presponsive to demand shifts driven by mothers' intended or realized labor supply (Manski, 1993). Similarly, provision

[12] Alternatively, birth thresholds for the eligibility of childcare subsidies or the admission to public childcare have been exploited (Gelbach (2002), Fitzpatrick (2010) and Bassok et al. (2014) for the U.S., Bauernschuster and Schlotter (2013) for Germany, and Berlinski et al. (2011) for Argentina).

[13] Felfe and Lalive (2013) who use the same instrument include state, but not county fixed effects. Bauernschuster and Schlotter (2013) exploit municipality variation in the supply of childcare due to cut-off rules; they include state, but not municipality (or county) fixed effects. Blau and Tekin (2007) use state dummies directly as instruments for the receipt of a childcare subsidy.

[14] This approach has been used in similar studies (Berlinski and Galiani, 2007; Blau and Tekin, 2007; Havnes and Mogstad, 2011) correcting standard errors for heteroskedasticity (Angrist and Krueger, 1999).

of subsidized childcare and mothers' labor supply in a given region might depend on the same macro-trends or shocks that are not represented in the control variables. These common influences might originate in economic, political, or cultural factors. Both specifications are more or less demanding regarding identifying assumptions in this respect (see sub-section 4.4 for discussion).

When these conditions are fulfilled, δ represents the Local Average Treatment Effect (LATE) of public childcare on employment for the sub-population of compliers (Imbens and Angrist, 1994). Mothers who decide to send their children to subsidized childcare as a result of an (exogenous) increase in local supply of childcare slots represent the complier population in our application.

Besides validity considerations, the regional variation in chidcare coverage must be a relevant instrument for the maternal childcare choice. We face weak instruments problems (Bound et al., 1995) in all specifications with covariates that become particularly severe in specification (b). Above all, this is a data issue, since the number of observations in the SOEP and the amount of variation is quite limited. We therefore exploit the German microcensus, a much larger data set providing considerably more (within-county) variation. Since the individual decision on a mode of childcare is not observed, we have to consider alternative estimation strategies.

4.2 Difference-in-differences

Without information on the individual childcare choice only an intention-to-treat (ITT) effect can be estimated. We can neither apply a classical difference-in-difference (DD) approach with clearly defined treatment/control groups and pre/post reform periods (e.g. Baker et al., 2008). Nor is it possible to use a generalized DD framework where certain regions are treated at various points in time (e.g. Cascio, 2009). As new laws on childcare were binding for all counties, we apply a DD design suggested by Havnes and Mogstad (2011).

This approach utilizes reform-induced temporal and spatial variation in the provision of childcare. Treatment and control groups are defined on the basis of the change in childcare coverage between post- and pre-reform levels ($\Delta_t cc_j = cc_{jt_1} - cc_{jt_0}$). The analysis is only based on those two points in time (t_0, t_1). Counties in the treatment (control) group have an above- (below-) median expansion of childcare. Due to restrictions on the childcare data (sub-section 3.1), the closest pre-reform period is $t_0 = 2002$. We use the latest available wave of the MC as post-reform period $t_1 = 2011$.[15] The regression specification of the DD model is as follows:

$$y_{ijt} = \alpha + \pi_1 Treat_{ij} + \pi_2 Post_t + \delta \left(Treat_{ij} * Post_t\right) + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \varepsilon_{ijt} \quad \text{(4a)}$$

$$y_{ijt} = \alpha + \pi_1 Treat_{ij} + \pi_2 Post_t + \delta \left(Treat_{ij} * Post_t\right) + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \mu_j + \varepsilon_{ijt} \quad \text{(4b)}$$

$Treat_{ij}$ denotes a dummy variable for the treatment status and $Post_t$ for the treatment period. Therefore unobserved differences in mothers employment outcomes between different years and the treatment and control group are controlled. The parameter δ of the interaction term $(Treat_{ij} * Post_t)$ measures the effect of interest. Under the assumptions that monthers' employment in treated and untreated counties would have developed similar without the expansion of childcare, δ represents the causal effect of the provision of childcare. Besides asymmetric shocks, there are several reasons for selection processes into the treatment and control groups (see sub-section 4.4 for details). In those instances mothers' employment in treated counties might evolve differently, even if there is no impact of subsidized childcare.

As before we distinguish two specifications: (4a) exploits differential variation between treatment and control group before and after the reform conditional on individual (X_{ijt}) and regional (X_{jt}) control variables. (4b) also conditions on regional fixed effects μ_j at the county level. This means that identification is solely based on within-county variation; the DD aproach only controls for different levels in childcare policies between treatment and control group through $Treat_{ij}$. As in the IV framework we estimate linear probability models for all binary outcome variables by OLS (Angrist, 2001).[16][17]

4.3 Regional fixed effects

Based on the regional expansion of childcare the division of treatment and control group is artificial and applying some threshold is arbitrary. We thus use a

[15] The main robustness analysis uses the years 2006 and 2011; this specification uses a base period t_0 after the introduction of the first reform.

[16] NOTE 1: Since we have continuous treatment/control assignments, we might refer to changes-in-changes literature, also check non-linear extension of DD.

[17] NOTE 2: See recent Pischke paper – might be better not to include 'controls' in DD regression, but rather as weights in balancing test.

more general ITT framework as an alternative. Identification is not only based on the differences between a two groups. Instead, quasi-experimental variation in the provision of childcare between all counties is exploited. We estimate the following equation by OLS:[18].

$$y_{ijt} = \alpha + \delta cc_{jt} + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \gamma_t + \varepsilon_{ijt} \tag{5a}$$

$$y_{ijt} = \alpha + \delta cc_{jt} + X'_{ijt}\beta_1 + X'_{jt}\beta_2 + \mu_j + \gamma_t + \varepsilon_{ijt} \tag{5b}$$

The gain in efficiency of using more information comes at the cost of stronger identifying assumptions. Asymmetric shocks or selection issues are relevant for each singly county, not only avarages between treatment and control groups. Therefore the difference between our two main specifications that either do not or do include fixed effects μ_j at the county level is more substantial compared to the DD model in terms of identification.

The underlying assumption of specification (5a) is that variation in childcare coverage cc_{jt} is exogenous conditional on observables X and a general time trend γ_t. Several (unobserved) mechanisms lead to a correlation between cc_{jt} and ε_{ijt}, though. The selection of childcare providers into certain counties with higher demand for childcare (or vice versa) is an example for such a relation. The observed cross-sectional variation in childcare coverage may thus be (in part) a result of the spatial matching between childcare providers and mothers with high labor market attachment. The effect of childcare on maternal employment would be biased when this part of variation is used for estimation. In contrast, the two-way fixed effects specification (5b) can be interpreted as a generalized difference-in-differences approach. The inclusion of county fixed effects controls for time-invariant unobserved factors that might be correlated with regional childcare provision. Identification is only based on within-county differences over time and therefore related to the quasi-experimental variation induced by policy reforms.[19]

We investigate subsample heterogeneity of the estimated effect by running separate estimations for single and married mothers, highly and poorly educated mothers and for mothers interviewed before and after the introduction of the *Elterngeld* in

[18] As mentioned the use of the linear probability model has been justified by Angrist (2001)

[19] This specification is similar to Bauernschuster et al. (2013) who estimate this model with aggregate data at the county level.

robustness analyses. In order to correct for possible serial correlation of the error terms, we cluster standard errors at the county-year level and additionally use non-parametric block bootstrapping (Betrand et al., 2004) as a robustness check for the clustered standard errors.

4.4 Identification

All of the aforementioned approaches and specifications are based on spatial and temporal variation in the publicly subsidized provision of childcare at the level of (West) German counties. There are several threats to identification in this setting (Felfe et al., 2014; Havnes and Mogstad, 2011):

(i) Macro-shocks might affect the treatment and control groups differently.

(ii) Childcare providers (parents) locating (migrating) to areas with high female labor force participation and a sufficiemt demand for (supply of) childcare may lead to a two-sided selection process.

(iii) Municipalities are interested in attracting qualified labor by offering or subsidizing childcare slots of sufficient magnitude and quality.

(iv) Parents are equipped with certain beliefs towards child-rearing and employment. They demand a certain amount of care and lobby or vote for local childcare policies according to their preferences.

(v) The new childcare laws explicitly call for a demand-oriented expansion of childcare. Regions that are doing better economically might face less shortages initially. Counties where excess demand is particularly high may thus initiate the largest expansion of childcare.[20]

(vi) The gradual increase in childcare availability opens the door for differential long-term trends in treatment and control counties.

Processes (i) through (iv) are more of general nature whereas (v) and (vi) relate to childcare policy reforms. This does not preclude the former to affect childcare expansion induced by reforms. All are related to two common problems in

[20] This is a problem in many of the studies based on regional variation in childcare expansion (see, e.g., Havnes and Mogstad, 2011 or Nollenberger and Rodríguez-Planas, 2011).

treatment/control setups: differential time trends unrelated to the treatment and compositional changes between those groups. Moreover, reverse causality plays an important role as childcare supply might adjust to demand. The crucial difference between specifications in all estimation strategies of this paper is whether or not fixed regional effects are controlled for. According to this distinction different parts of the variation in childcare coverage are exploited for identification. All of the listed problems apply unconditionally to cross-sectional anlyses. Assumptions with respect to the temporal within-county part are less demanding.

Childcare expansion is assumed to be exogenous conditional on a number of intervening variables. Covariates in the estimations (sub-section 4.5) are supposed to control for several of the mechanisms depicted above. Mother- and household-specific characteristics (e.g. marital status, the number and age of children, other household income as well as age, qualification and labor market experience of the mother) reflect heterogeneity in preferences, financial incentives and capabilities in terms of mothers' labor market participation and utilization of childcare. These variables control primarily for compositional changes across the treatment and control units which might lead to shifts in maternal labor supply and demand for public childcare. Regional variables (population density, gross domestic product, female employment rate, fertility) approximate structural differences between counties. These might account for systematic differences in the demand for and supply of childcare which could lead to differential trends between treatment and control groups.

We argue that conditional on these covariates, the variation emanating from childcare reforms can be considered exogenous. This is a result of the implementation process in Germany (sub-section 2.3). Implementation involves a lengthy process at different administrative levels that consists of planning and projection of demand, applications for state-funding filed by local providers, and approval of proposal by state authorities (Felfe and Lalive, 2013):

(1) There is substantial error in local projections of childcare demand which has been documented (Hüsken, 2010, 2011). Planning is organized at the local level and those errors are not evenly distributed.

(2) Municipalities are capacity-constrained in terms of financial scope, qualified

personnel, or suitable construction grounds (BMFSFJ, 2011, 2012, 2013). Targets are therefore rarely met in the projected time frame.

(3) There are often considerable delays in the approval within the state administration.

As a result of severe shortages in supply, childcare providers operate with waiting lists. Families who sign up their children early are given prefered access. Children with single or working parents and with siblings may jump waiting lists. These different reasons generate exogenous variation between the municipalities which we observe at the county level and exploit for identification. These arguments only hold for the reform-induced expansion of childcare. The DD aproach as well as IV and panel specifications with county fixed effects explicitly rely on this part of the variation.

Descriptive evidence illustrates the substantial increase in childcare coverage for children aged under 3. The average coverage rate has increased monotonically between 2006 (2007)[21] and 2011 in West Germany (Tab. A1 in the Appendix). Overall (full-time) coverage has almost tripled from 7% (2.5%) to almost 20% (6.5%). This poses a marked supply shock, i.e. a treatment of significant magnitude across West Germany. An expansion of publicly subsidized childcare for children under the age of 3 started already at the beginning of the 2000s (Fig. A1 in the Appendix), but the tempo increased considerably in the middle ot the 2000s. The monotonic increase holds for each single state which demonstrates that compliance has been comprehensive.

Regional heterogeneity has been reduced in relative terms as measured by the Theil index during this period of expansion (bottom of Tab. A1).[22] Not only have childcare slots become more equally distributed across all of West Germany, but also between and within federal states. Inequality in childcare provision has decreased more within than between states. This holds for overall and full-time coverage. As of 2011 there is still considerable regional variation between and within West German states in the provision of childcare, considerably more so for full-time slots.

[21] We only have data for full-time slots from 2007 onwards.

[22] Bauernschuster et al. (2013) argue that heterogeneity increased as measured by standard deviations. Yet, this is rather a mechanical effect depending on the level of childcare coverage. This is a question of relative and absolute heterogeneity; it is not a priori clear what is more relevant for identification.

It is hard to pin down empirical evidence that the implementation of reforms generated idiosyncratic variation in the provision of childcare. A detailed visualization of how the spatial distribution of childcare coverage has evovled over the post-reform period provides some guidance in that regard (Fig. A3 in the Appendix). The considerable within- and between-state variation in the cross-section, but also over time is confirmed. In 2006 certain counties start from a much higher level than others. In 2006 there are several regional clusters with high coverage, e.g. the north of Bavaria or the south of Rhineland Palatinate, but also large cities as Hamburg or Munich. More importantly the expansion does not proceed with uniform tempo. Certain counties move faster than others that catch up the following year or later. Although many of the aforementioned focal points keep their edge, it is also visible that many counties catch up to a certain degree over time.

Comparing the development of overall and full-time coverage also reveals some important insights. Some of the counties/regions with above-average overall coverage also provide a high number of full-time slots (some large cities as Hamburg or Munich), some do not. The Bavarian north-south divide does not exist for full-time care. Some counties/regions that proceeded more quickly in expanding overall childcare coverage also have invested more in full-time slots. On the other hand, certain counties, in some instances regions or whole states (e.g. North-Rhine Westphalia), which have long lagged behind in overall coverage, moved to the top in terms of full-time care. These findings underline the erratic spatial pattern during the expansion and provide evidence for exogenous variation in chilcare supply.

4.5 Variables and sample

Dependent variables

The effect of subsidized childcare for the group of mothers with young children will likely be heterogeneous for different margins of labor supply (sub-section 2.1). We therefore estimate the relationship for various outcome variables y_{ijt} capturing the extensive or intensive margin in each of the different empirical strategies:

(1) *Participation*: $y_{ijt}^{(1)}$ is a dummy variable where $y_{ijt}^{(1)} = 1$, if the mothers hours of work are positive, i.e. $h_{ijt} > 0$, and $y_{ijt}^{(1)} = 0$ otherwise. This is an overall indicator for mothers' labor supply at the extensive margin.

(2) *Quality participation*: $y_{ijt}^{(2)}$ is a dummy variable where $y_{ijt}^{(2)} = 1$, if $h_{ijt} > 0$ and the mother has a regular employment contract (as observed in the data). Otherwise $y_{ijt}^{(2)} = 0$. This means she does not work in a so-called mini-job without income tax and reduced social security contributions. This measure is supposed to capture a mother's 'true' labor market integration. $y_{ijt}^{(2)}$ could be more closely related to childcare coverage when marginal employment can be realized without having the child in formal care.

(3) *Full-time participation*: $y_{ijt}^{(3)}$ is a dummy variable where $y_{ijt}^{(3)} = 1$, if $h_{ijt} > 0$ and the self-assessed employment status is full-time; $y_{ijt}^{(3)} = 0$ otherwise. This outcome measures the influence of public childcare on increasing the full-time share among empoyed women.

(4) *Marginal employment*: $y_{ijt}^{(4)}$ is a dummy variable where $y_{ijt}^{(4)} = 1$, if the mother works in a mini-job; the state $y_{ijt}^{(4)} = 0$ includes quality participation ($h_{ijt} > 0$) and non-employment ($h_{ijt} = 0$). When marginal employment does not necessarilydepend on public childcare, one would expect a different effect compared to other participation measures.

(5) *Hours of work*: $y_{ijt}^{(5)}$ is a cardinal variable whereas $y_{ijt}^{(5)} = h_{ijt}$, if $h_{ijt} > 0$. It measures the intensive margin of labor supply and is only observed, if $y_{ijt}^{(1)}$. Since we can only analyze the intensive labor supply margin for women participating in the labor market, we adjust for all hours estimates for a potential selection bias.

We follow standard practice and apply a Heckit two-step model (Heckman (1979)). Under the assumption that the error terms in the respective structural and selection equations are distributed jointly normal, we add the inverse Mills ratio λ_{ijt} in each of our structural equations as a selection correction. The selection equation is estimated for the respective pooled samples and includes the dummies for being a lone mother and maried as well as other household income as exclusion restrictions.

Explanatory variables

The explanatory variable of interest measures the provision of childcare for children aged under 3 at the county level. In the majority of specifications for the period

2006 to 2011 the child care coverage rate cc_{jt} is available defined as the percentage of children of this age group using subsidized formal child care county j in year t. Between 2007 and 2011 we have information on the split of full-time cc_{jt}^{FT} and part-time coverage rates cc_{jt}^{PT} with $cc_{jt} = cc_{jt}^{FT} + cc_{jt}^{PT}$. The theoretical considerations have shown that the mother's employment decision will depend on the quantity and quality of different childcare options available to her. It is conceiveable that she prefers a certain number of working hours that can only be reached when the child is in full-time childcare.

We therefore use the following flexible specification for the IV and reduced-form estimations:

$$\kappa cc_{jt} = \kappa_1 cc_{jt}^{FT} + \kappa_2 cc_{jt}^{PT} \tag{6a}$$

$$\delta cc_{jt} = \delta_1 cc_{jt}^{FT} + \delta_2 cc_{jt}^{PT} \tag{6b}$$

We use overall childcare coverage cc_{jt} and full-time coverage cc_{jt}^{FT} for robustness analyses. We define the treatment and control groups in the standard DD model based on overall coverage cc_{jt} and use a distinction on the basis of only full-time coverage cc_{jt}^{FT} as a robustness check. Descriptive statistics on childcare coverage are documented in Tab. A1 in the Appendix.

The general set of individual control variables for the mother includes her age (included in linear and quadratic form in all specifications), the level of qualification, her marial status and nationality. For the estimations based on the SOEP we also exploit information on the mother's labor market experience as well as additional household information. We distinguish single and cohabiting mothers, the gender of the (youngest) child, and other household income. These variables account for heterogeneity in mother'S preferences, variation in financial incentives and capabilities determining her labor market participation and utilization of childcare.

Hüsken (2010) runs various regressions to detect determinants of regional differences in child care coverage for under three year olds based on data from 2010. She finds that the degree of urbanicity, gdp per capita, the female employment rate, the proportion of employed women working part time and the proportion of highly skilled workers in the area are positively correlated with child care coverage for children under the age of three in West Germany, while the latter is negatively correlated with the regional fertility and unemployment rate. Therefore we control for the endogeneity of childcare supply by including these variables measured at the county level as control variables. We will conduct a similar test regression for our sample.

Sample definition

We analyze the effect of subsidized childcare for children aged under 3 on their mothers' labor supply. In this version of the paper we focus on West Germany where the childcare expansion was largest. Due to administrative reforms that changed the alignment of counties, we do not have consistent panel data on childcare coverage for East Germany. There are further restrictions concerning the childcare data (sub-section 3.1). Information on full-time coverage is only available between 2007 and 2011. We therefore restrict our main analysis to this period.

Robustness analyses are conducted for a number of sub-groups for this time period. The validity of robustness analyses with longer time periods is limited, though. Before 2006 regional childcare data are only available for the years 1998 and 2002. For those years childcare statistics are based on available slots and not the actual utilization (sub-section 3.3). Changes in the survey design of the MC that affected the labor supply measures further diminishes comparability. We present some robustness checks for 2006 to 2011 and 2002 to 2011 bearing those limitations in mind.[23]

[23] ADD table on sample sizes for SOEP and MC that demonstrates the size advantge within counties.

5 Results

5.1 Instrumental variables

IV estimates for the years 2007 to 2011 are based on SOEP data. We use the full-time and part-time childcare coverage as excluded instruments for the individual utilization of childcare. We compare OLS estimates with the endogenous childcare decision as explanatory variable with 2SLS estimates. Three specifications are distinguished for both approaches: (1) only childcare indicators cc_{jt}^{FT}, cc_{jt}^{PT} without covariates; (2) adding observables X and year fixed effects γ_t; (3) adding county fixed effects μ_j (Tab. 1).

Table 1: IV estimates, effects of childcare provision on labor supply

	OLS			**IV**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Extensive margin						
Participation						
	0.383***	0.305***	0.286***	0.388*	0.347	-0.119
	(0.032)	(0.034)	(0.043)	(0.162)	(0.301)	(0.432)
Quality participation						
	0.391***	0.322***	0.320***	0.533***	0.689*	0.210
	(0.033)	(0.034)	(0.041)	(0.140)	(0.308)	(0.385)
Full-time employment						
	0.097***	0.089***	0.091**	0.227*	0.342	0.060
	(0.022)	(0.024)	(0.030)	(0.096)	(0.211)	(0.223)
Marginal employment						
	-0.008	-0.017	-0.034	-0.145	-0.341	-0.329
	(0.022)	(0.026)	(0.033)	(0.097)	(0.245)	(0.313)
Weak IV tests						
Partial R-sq.				0.041	0.013	0.006
Robust F				15.869	5.530	2.573
N	1159	1159	1159	1159	1159	1159
Intensive margin						
Hours worked						
	4.820***	4.725***	4.848*	17.915**	22.003**	7.183
	(1.268)	(1.355)	(2.060)	(5.595)	(7.869)	(12.676)
Weak IV tests						
Partial R-sq.				0.068	0.043	0.009
Robust F				22.430	11.387	1.078
N	511	511	511	511	511	511

Notes: Covariates in different specifications: (1) CC provision only; (2) CC provision, covariates at individual, household, county level, year fixed effects, (2) CC provision, covariates at individual, household, county level, year & county fixed effects; standard errors in parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: SOEP, own calculations.

The OLS models yield significantly positive findings for all outcomes except for marginal employment. Coefficients are negative, albeit not statistically significant. As indicated above highly significant coefficients might reflect the endogenous nature of the childcare choice. Therefore instrumental variables are needed. Looking at specifications (1) and (2) of the IV model, the OLS findings seem to be corroborated as the effects remain positive and significant. The magnitude of IV coefficients is larger compared to OLS. Theory would rather predict an upward bias because the choices of childcare and labor supply might, e.g., both be related to unobserved traits. While the instrument is relevant in (1), it loses power and remains only reliably strong for the intensive margin estimate in (2) (weak IV tests at bottom of Tab. 1).

Specification (3) is solely based on quasi-experimental within-county variation in the expansion of subsidized care. Adding county fixed effects kills identification in the first stage, however. The instruments fail the weak IV tests thoroughly (Tab. 1). The effect of childcare utilization becomes insignificant in the second stage of the IV model. Given the size of the SOEP data within-county variation in childcare coverage seems to be too small to serve as a reliable instrument. It is therefore impossible to check the robustness of specification (2) in that regard on the basis of the SOEP.

There is an obvious trade-off between stricter assumptions and sufficient variation to identify the effect of interest. This uncertainty is worth noting given that previous studies rely on those stricter assumptions. We therefore turn to the MC data that promise to offer a larger amount of identifying variation. The estimates will not be directly comparable, however, we are only able to identify an ITT effect (as opposed to the LATE with IV) due to lack of information on the households' childcare choice.

5.2 Difference-in-differences

As decribed the DD approach is based only on two waves of data. The assignment of treatment and control group is based on overall childcare coverage. Due to several data restrictions we consider to samples. Our main specification with consistent childcare and employment data includes the years 2006 and 2011. A robustness check is provided based on the pre-reform year 2002 and 2011 bearing potential data inconsistencies in mind.

The DD model as such is similar to the IV specification with regional regional fixed effects as it exploits only the temporal differences between the treatment and the control group. We look first at SOEP estimates for comparative reasons and then turn to the results based on the MC. We nevertheless compare specifications with and without county fixed effects.

SOEP results

We do not get any significant treatment effect for any of the outcomes, in any of the specifications, or for any time period (Tab. 3). This confirms the IV findings that are based on within-region variation and reflecs the limited amount of identifying information in the SOEP data.

Table 2: DD estimates, effects of childcare provision on labor supply

	2006, 2011			**2002, 2011**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Extensive margin						
Participation						
Treatment	-0.058	-0.069	0.089	0.069	0.018	0.106
	(0.098)	(0.098)	(0.149)	(0.081)	(0.084)	(0.129)
Quality participation						
Treatment	-0.042	-0.068	-0.053	0.104	0.028	0.001
	(0.089)	(0.086)	(0.123)	(0.080)	(0.078)	(0.108)
Full-time employment						
Treatment	0.042	0.035	0.057	-0.042	-0.062	-0.067
	(0.049)	(0.048)	(0.085)	(0.042)	(0.042)	(0.065)
Marginal employment						
Treatment	-0.017	-0.001	0.142	-0.034	-0.010	0.105
	(0.062)	(0.063)	(0.117)	(0.058)	(0.056)	(0.089)
N	398	398	398	494	494	494
Intensive margin						
Hours worked						
Treatment	4.085	4.056	-4.408	0.685	0.467	-5.060
	(3.841)	(3.586)	(10.044)	(3.698)	(3.315)	(9.810)
N	173	173	173	181	181	181

Notes: Covariates in different specifications: (1) CC provision only; (2) CC provision, covariates at individual, household, county level, year fixed effects, (2) CC provision, covariates at individual, household, county level, year & county fixed effects; standard errors in parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: SOEP, own calculations.

MC results

As for the estimates based on the SOEP the coefficients representing the treatment effect turn out to be highly insignificant for all outcomes, throughout all specifications and for both sample periods (Tab. 3). Note that the size of the MC exceeds the SOEP sample size considerable, e.g. for the extensive margin between 2006 and 2011 by the factor 30.

Table 3: DD estimates, effects of childcare provision on labor supply

	2006, 2011			**2002, 2011**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Extensive margin						
Participation						
Treatment	0.013	0.012	0.010	-0.009	0.003	0.027
	(0.018)	(0.016)	(0.023)	(0.020)	(0.019)	(0.038)
Quality participation						
Treatment	0.007	0.007	0.006	-0.012	0.001	0.031
	(0.020)	(0.017)	(0.025)	(0.019)	(0.018)	(0.034)
Full-time employment						
Treatment	0.014	0.017	0.031	0.006	0.016	0.014
	(0.015)	(0.014)	(0.024)	(0.018)	(0.018)	(0.031)
Marginal employment						
Treatment	-0.005	-0.006	-0.014	-0.001	-0.001	-0.005
	(0.012)	(0.011)	(0.017)	(0.012)	(0.012)	(0.021)
N	12199	12199	12199	13349	13349	13349
Intensive margin						
Hours worked						
Treatment	-0.272	0.065	0.118	0.196	0.330	-0.407
	(0.521)	(0.527)	(0.548)	(0.699)	(0.693)	(0.734)
N	9514	9514	9514	8933	8933	8933

Notes: Covariates in different specifications: (1) CC provision only; (2) CC provision, covariates at individual, household, county level, year fixed effects, (2) CC provision, covariates at individual, household, county level, year & county fixed effects; standard errors in parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: MC, own calculations.

Due to the limited availability of data for the pre-reform years it is not easy to visualize trends of important variables for the treatment and control groups. [Graphs will be provided in the next version of this paper.] It is not clear whether this is a substantive finding: It could also be explained by a lack of identifying variation or could be driven by the violation of the common trend assumption.

5.3 Regional fixed effects

The underlying idea of this approach ist to exploit all of the within-county variation in childcare expansion. The DD model uses only temporal difference between two groups of counties and two waves of data. Given the exogeneity of this information, the regression with fixed county effects is more efficient as it not only uses more within-variation, but also several years of data. Our baseline period is 2006 to 2011. We can differentiate between full-time and part-time childcare coverage. Again, we use the SOEP estimates as a benchmark for comparison with the other approaches. Our main estimates are then based on the MC data.

SOEP results

Contrary to the insignificant DD coefficients, we see that the pattern of coefficients in large parts replicates the IV results. We find positive effects on all outcome variables except for marginal employment in specifications (1) through (3) without regional fixed effects (Tab. 4). Only some of them turn out to be significant. This might reflect the efficiency loss in those reduced-form estimates in comparison to the more efficient IV estimates. Coefficients for marginal employment are negative, albeit not statistically significant.

Specification (4) includes county fixed effects. Coefficients are in general smaller in magnitude and very imprecisely estimated. Again this reflects the limnited variation for identification. It might also hint at upwardly biased childcare effects in specifications (1) through (3).

Table 4: Regression estimates, effects of childcare provision on labor supply

	(1)	(2)	(3)	(4)
Extensive margin				
Participation				
Full-time childcare	0.634	0.192	0.410	0.164
	(0.460)	(0.518)	(0.594)	(2.468)
Part-time childcare	0.602	0.362	0.546	-0.472
	(0.416)	(0.440)	(0.470)	(1.481)
Quality participation				
Full-time childcare	1.217**	0.963*	1.265*	0.719
	(0.394)	(0.444)	(0.503)	(2.364)
Part-time childcare	0.385	0.276	0.497	0.515
	(0.381)	(0.451)	(0.469)	(1.395)
Full-time employment				
Full-time childcare	0.354	0.411	0.515	-0.615
	(0.241)	(0.304)	(0.343)	(1.330)
Part-time childcare	0.392	0.263	0.300	0.134
	(0.230)	(0.275)	(0.299)	(0.795)
Marginal employment				
Full-time childcare	-0.583*	-0.771*	-0.855*	-0.555
	(0.263)	(0.334)	(0.392)	(1.521)
Part-time childcare	0.217	0.087	0.049	-0.987
	(0.265)	(0.318)	(0.340)	(0.970)
N	1180	1180	1180	1180
Intensive margin				
Hours worked				
Full-time childcare	52.317**	68.158***	71.535**	26.829
	(16.340)	(20.228)	(23.450)	(105.119)
Part-time childcare	29.096	27.747	28.547	39.772
	(16.339)	(19.053)	(20.833)	(78.281)
N	517	517	517	517

Notes: Covariates in different specifications: (1) CC provision only, (2) CC provision, covariates at individual, household, county level, (3) CC provision, covariates at individual, household, county level, year fixed effects, (4) CC provision, covariates at individual, household, county level, year & county fixed effects; standard errors in parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: SOEP, own calculations.

Microcensus results

The regression estimates based on the MC yield patterns that are comparable to the SOEP findings (Tab. 5). The coefficients are smaller in magnitude and in general more precisely estimated. We find mostly significant findings in specifications (1) to (3) that do not contain county fixed effects. The added variation seems to support identification. Marginal employment is again the exception with mostly negative coefficients.

The main finding here is that the coefficients for the extensive margin outcomes

Table 5: Regression estimates, effects of childcare provision on labor supply

	(1)	(2)	(3)	(4)
Extensive margin				
Participation				
Full-time childcare	0.192**	0.328***	0.276**	0.057
	(0.071)	(0.086)	(0.093)	(0.299)
Part-time childcare	0.308***	0.266**	0.232*	0.214
	(0.078)	(0.090)	(0.095)	(0.229)
Quality participation				
Full-time childcare	0.372***	0.514***	0.210*	-0.155
	(0.069)	(0.092)	(0.099)	(0.292)
Part-time childcare	0.574***	0.476***	0.257**	0.408
	(0.085)	(0.100)	(0.098)	(0.232)
Full-time employment				
Full-time childcare	0.305***	0.118	0.156	0.190
	(0.080)	(0.076)	(0.084)	(0.306)
Part-time childcare	0.075	0.032	0.065	0.016
	(0.063)	(0.070)	(0.076)	(0.203)
Marginal employment				
Full-time childcare	-0.313***	-0.255***	0.036	0.247
	(0.055)	(0.058)	(0.048)	(0.231)
Part-time childcare	-0.385***	-0.288***	-0.071	-0.352*
	(0.047)	(0.048)	(0.042)	(0.166)
N	32893	32893	32893	32893
Intensive margin				
Hours worked				
Full-time childcare	14.160**	6.019	12.632**	22.103*
	(4.277)	(3.568)	(4.362)	(10.513)
Part-time childcare	-3.214	-2.588	1.991	8.038
	(2.495)	(2.876)	(3.249)	(8.873)
N	26547	26547	26547	26547

Notes: Covariates in different specifications: (1) CC provision only, (2) CC provision, covariates at individual, household, county level, (3) CC provision, covariates at individual, household, county level, year fixed effects, (4) CC provision, covariates at individual, household, county level, year & county fixed effects; standard errors in parentheses; *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.
Source: Microcensus, own calculations.

become markedly smaller and turn insignificant when county fixed effects are included. In spite of the relatively large amount of variation we do not get significant effects. We interpret this as evidence that participation effects of the expansion of childcare are smaller than previously thought, potentially even absent.

Yet, we do find a marginally statistically significant effect for full-time coverage for the working hours of employed mothers. We find that one more child care slot per 100 children increases the average working hours of already employed mothers by roughly 15 minutes. The point estimate for the regression model with county fixed effects on SOEP data was of similar magnitude, but very imprecisely estimated.

A very cautious interpretation might be that already employed mothers with small children extended their working hours as childcare coverage, in particular the supply of full-time places, improved.

6 Discussion and conclusion

We have estimated the effect of the expansion of subsidized childcare on the labor supply of mothers with children aged 1 to 3 in Germany. We have found significantly positive effects of subsidized childcare on maternal labor supply based on the overall variation in childcare provision. This holds for all measures of labor supply at the extensive and intensive margin as well as IV and reduced-form frameworks. Results change dramatically when identification is restricted to quasi-experimental within-county variation. Childcare coverage indicators turned out to be weak instruments in estimations with the SOEP. The difference-in-difference framework which exploits the same type of variation also yields insignificant estimates for all outcome variables. Using the larger MC data with more within-county variation and employing a more efficient regression approach with county fixed effects we find marginally significant effects at the extensive margin.

In line with previous assessments (Blau, 2003; Blau and Currie, 2006) we conclude that estimates based on quasi-experimental variation are very sensitive with respect to identifying assumptions and related model specifications. This has far-reaching consequences for the substantive implications of the findings. It casts doubt on previous estimates of significant participation effects – most of them did neither rely exclusively on quasi-experimental variation, nor provide robustness checks with county fixed effects.

Taking the findings based on the arguably more credible exogenous quasi-experimental variation at face value there is no identifiable extensive margin effect of childcare on labor supply of mothers. However, already employed mothers may increase their working hours. A marginal increase of child care availability (one additional full-time child care slot per hundred children) increases average working hours of already working mothers by roughly 15 minutes per hour. Our results are in line with comparable studies for Sweden and Norway that found similar patterns. They also confirm the small extensive margin elasticities found in structural approaches.

Thus, our paper contributes to the existing empirical literature in several ways. First, we provide evidence for the impact of a large child care expansion in Germany - a country with very low initial child care coverage and very low maternal employment. Second, in contrast to most of the international literature that focusses on children aged three years and older, we consider children in the age group of 1 to 3 years. Second, we distinguish between the extensive and the intensive margin, which appears to be crucial since we find heterogenous effects for the two margins. Moreover, we distiguish between the availability of part-time and full-time child care which has very different implications for the employment opportunities of mothers.

References

Afentakis, Anja and Wolf Bihler, "Das Hochrechnungsverfahren beim unterjährigen Mikrozensus ab 2005," *Wirtschaft und Statistik*, 2010, *10/2010*, 1039–1048.

Anderson, Patricia M. and Phillip B. Levine, "Child Care and Mothers' Employment Decisions," in David E. Card and Rebecca M. Blank, eds., *Finding Jobs: Work and Welfare Reform*, Russel Sage Foundation, 2000, pp. 420–462.

Angrist, Joshua D., "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice," *Journal of Business & Economic Statistics*, 2001, *19* (1), 2–16.

_ **and Alan B. Krueger**, "Empirical Strategies in Labor Economics," in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam et al.: Elsevier, 1999, pp. 1277–1366.

Averett, Susan L., H. Elizabeth Peters, and Donald M. Waldman, "Tax Credits, Labor Supply, And Child Care," *The Review of Economics and Statistics*, February 1997, *79* (1), 125–135.

Baker, Michael, Jonathan Gruber, and Kevin Milligan, "Universal childcare, maternal labor supply and family well-being," *Journal of Political Economy*, 2008, *116* (4), 709–745.

Bassok, Daphna, Maria Fitzpatrick, and Susanna Loeb, "Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia," *Journal of Urban Economics*, 2014, *83* (0), 18–33.

Bauernschuster, Stefan and Martin Schlotter, "Public Child Care and Mothers' Labor Supply – Evidence from Two Quasi-Experiments," Technical Report, CESifo Working Paper No. 4191, Ifo Institute – Leibniz Institute for Economic Research at the University of Munich, Munich 2013.

_ **, Timo Hener, and Helmut Rainer**, "Does the Expansion of Public Child Care Increase Birth Rates? Evidence from a Low-Fertility Country," Technical

Report, ifo Working Paper No. 158, Ifo Institute – Leibniz Institute for Economic Research at the University of Munich, Munich 2013.

Becker, Gary S., *A Treatise on the Family* number beck81-1. In 'NBER Books.', National Bureau of Economic Research, Inc, 1981.

Berger, Mark C. and Dan A. Black, "Child Care Subsidies, Quality of Care, and the Labor Supply of Low-Income, Single Mothers," *The Review of Economics and Statistics*, 1992, *74* (4), 635–642.

Berlinski, Samuel and Sebastian Galiani, "The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment," *Labour Economics*, 2007, *14*, 665–680.

_ , _ , and Patrick J. McEwan, "Preschool and Maternal Labor Market Outcomes: Evidence from a Regression Discontinuity Design," *Economic Development and Cultural Change*, 2011, *59* (2), 313–344.

Betrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust difference-in-difference estimates," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Bick, Alexander, "The quantitative role of child care for female labor force participation and fertility," Technical Report, MPRA PaperNo. 25474, University Library of Munich, Germany August 2011.

Blau, David, *The Child Care Problem*, New York: Russel Sage Foundation, 2001.

_ , "Child Care Subsidy Programs," in "Means-Tested Transfer Programs in the United States" NBER Chapters, National Bureau of Economic Research, Inc, Jan-Jun 2003, pp. 443–516.

_ **and Erdal Tekin**, "The determinants and consequences of child care subsidies for single mothers in the USA," *Journal of Population Economics*, 2007, *20* (4), 719–741.

_ **and Janet Currie**, "Pre-School, Day Care, and After-School Care: Who's Minding the Kids?," in Erik Hanushek and F. Welch, eds., *Erik Hanushek and*

F. Welch, eds., Vol. 2 of *Handbook of the Economics of Education*, Elsevier, June 2006, chapter 20, pp. 1163–1278.

Blau, David M. and Alison P. Hagy, "The Demand for Quality in Child Care," *Journal of Political Economy*, February 1998, *106* (1), 104–146.

Blau, David M and Philip K Robins, "Child-Care Costs and Family Labor Supply," *The Review of Economics and Statistics*, August 1988, *70* (3), 374–81.

Blau, David M. and Philip K. Robins, "Child Care Demand and Labor Supply of Young Mothers over Time," *Demography*, 1991, *28* (3), 333–351.

Blundell, Richard, Allen Duncan, Julian MacCrae, and Costas Meghir, "Household Labour Supply, Childcare Costs and Tax Credits," Technical Report 2000.

_ , _ , Julian McCrae, and Costas Meghir, "The Labour Market Impact of the Working Families' Tax Credit," *Fiscal Studies*, 2000, *21* (1), 75–104.

BMFSFJ, "Zweiter Zwischenbericht zur Evaluation des Kinderförderungsgesetzes.," Technical Report, Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ), Bericht der Bundesregierung für das Berichtsjahr 2010 2011.

_ , "Dritter Zwischenbericht zur Evaluation des Kinderförderungsgesetzes.," Technical Report, Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ), Bericht der Bundesregierung für das Berichtsjahr 2011 2012.

_ , "Vierter Zwischenbericht zur Evaluation des Kinderförderungsgesetzes.," Technical Report, Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ), Bericht der Bundesregierung für das Berichtsjahr 2012 2013.

Boca, D. Del and D. Vuri, "The mismatch between employment and child care in Italy: The impact of rationing," *Journal of Population Economics*, 2007, *20* (4), 805–832.

Boca, Daniela Del, "The Effect of Child Care and Part Time Opportunities on Participation and Fertility Decisions in Italy," *Journal of Population Economics*, 2002, *15* (3), 549–573.

_ **and Robert M. Sauer**, "Life cycle employment and fertility across institutional environments," *European Economic Review*, 2009, *53*, 274–292.

Bound, John, David A. Jaeger, and Regina M. Baker, "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 1995, *90* (430), 443–450.

Büchel, Felix and C. Katharina Spie"s, "Kindertageseinrichtungen und Müttererwerbstätigkeit: neue Erkenntnisse zu einem bekannten Zusammenhang," *Vierteljahrshefte zur Wirtschaftsforschung*, 2002, *71* (1), 95–113.

Cascio, Elizabeth U., "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools," *The Journal of Human Resources*, 2009, *44* (1), 140–170.

Choné, Philippe, David le Blanc, and Isabelle Robert-Bobée, "Female Labor Supply and Child Care in France," Technical Report, CESifo Working Paper No. 1059, Center for Economic Studies & Ifo Institute for Economic Research 2003.

Coneus, Katja, Kathrin Goeggely, and Grit Muehler, "Maternal employment and child care decision," *Oxford Economic Papers*, 2008, *61*, i172–i188.

Connelly, Rachel, "The Effect of Child Care Costs on Married Women's Labor Force Participation," *The Review of Economics and Statistics*, February 1992, *74* (1), 83–90.

_ **and Jean Kimmel**, "The Effect of Child Care Costs on the Employment and Welfare Recipiency of Single Mothers," *Southern Economic Journal*, January 2003, *69* (3), 498–519.

Cowell, Frank A., *Measuring Inequality*, Oxford, UK: Oxford University Press, 2000.

Duncan, Alan and Christopher Giles, "Labour Supply Incentives and Recent Family Credit Reforms," *Economic Journal*, January 1996, *106* (434), 142–55.

— , **Gillian Paull, and Jayne Taylor**, "Mothers' employment and the use of childcare in the UK," Technical Report, IFS Working Papers No. W01/23, Institute for Fiscal Studies (IFS) October 2001.

Felfe, Christina and Rafael Lalive, "Early Child Care and Child Development: For Whom It Works and Why," Technical Report, SOEPpapers on Multidisciplinary Panel Data Research No. 536, German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung, DIW) 2013.

— , **Michael Lechner, and Petra Thiemann**, "After-school care and parents' labor supply," Technical Report, Swiss Institute for International Economics and Applied Economics Research (SIAW), University of St. Gallen 2014.

Fitzpatrick, Maria Donovan, "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten," *Journal of Labor Economics*, 01 2010, *28* (1), 51–85.

— , "Revising Our Thinking About the Relationship Between Maternal Labor Supply and Preschool," *Journal of Human Resources*, 2012, *47* (3), 583–612.

Gathmann, Christina and Bjoern Sass, "Taxing Childcare and its Effects on Children and their Families in Germany," Technical Report, SOEPpapers on Multidisciplinary Panel Data Research No. 438, German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung, DIW) 2012.

Gelbach, Jonah B., "Public Schooling for Young Children and Maternal Labor Supply," *American Economic Review*, 2002, *92* (1), 307–322.

Gennetian, Lisa A., Anna Gassman-Pines, Aletha C. Huston, Danielle A. Crosby, Young Eun Chang, and Edward D. Lowe, "A Review of Child Care Policies in Experimental Welfare and Employment Programs," Technical Report, MDRC Working Paper Series No. 1, Manpower Demonstration Research Corporation (MDRC), New York 2001.

Geyer, Johannes, Peter Haan, and Katharina Wrohlich, "The Effects of Family Policy on Mothers' Labor Supply. Combining Evidence from a Structural

Model and a Natural Experiment," Technical Report, SOEPpapers on Multidisciplinary Panel Data Research No. 645, German Institute for Economic Research (DIW Berlin), Berlin 2014.

Goux, Dominique and Eric Maurin, "Public school availability for two-year olds and mothers' labour supply," *Labour Economics*, December 2010, *17* (6), 951–962.

Gustafsson, Siv and Frank Stafford, "Child Care Subsidies and Labor Supply in Sweden," *Journal of Human Resources*, 1992, *27* (1), 204–230.

Haan, Peter and Katharina Wrohlich, "Can child care policy encourage employment and fertility? Evidence from a structural model," *Labour Ecnomics*, 2011, *18*, 498–512.

Han, Wenjui and Jane Waldfogel, "Child Care Costs and Women's Employment: A Comparison of Single and Married Mothers With Pre-School-Aged Children," *Social Science Quarterly*, 2001, *82* (3), 552–568.

Havnes, Tarjei and Magne Mogstad, "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 2011, *95* (11), 1455–1465.

Heckman, James J., "Effects of Child-Care Programs on Women's Work Effort," *Journal of Political Economy*, 1974, *82* (2), pp. S136–S163.

_ , "Sample Selection Bias as a Specification Error," *Econometrica*, 1979, *47* (1), 153–162.

Helmcke, Thomas, "Regionalstatistik auf europäischer und nationaler Ebene," *Wirtschaft und Statistik*, 2008, *3*, 207–216.

Hüsken, Katrin, *Kita vor Ort. Betreuungsatlas auf Ebene der Jugendamtsbezirke 2008* 2010.

_ , *Kita vor Ort. Betreuungsatlas auf Ebene der Jugendamtsbezirke 2010* 2011.

II, Charles L. Baum, "A Dynamic Analysis of the Effect of Child Care Costs on the Work Decisions of Low-Income Mothers with Infants," *Demography*, 2002, *39* (1), 139–164.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Immervoll, Herwig and David Barber, "Can Parents Afford to Work? Childcare Costs, Tax-Benefit Policies and Work Incentives," Technical Report, IZA Discussion Paper No. 1932, Institute for the Study of Labor (Forschungsinstitut zur Zukunft der Arbeit, IZA) January 2006.

Jacobi, Lena and Jochen Kluve, "Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany," *Zeitschrift für ArbeitsmarktForschung*, 2007, *40* (1), 45–64.

Kimmel, Jean, "The Effectiveness of Child-Care Subsidies in Encouraging the Welfare-to-Work Transition of Low-Income Single Mothers," *American Economic Review*, May 1995, *85* (2), 271–75.

_ , "Child Care Costs As A Barrier To Employment For Single And Married Mothers," *The Review of Economics and Statistics*, May 1998, *80* (2), 287–299.

Kreyenfeld, Michaela and Karsten Hank, "Does the availability of child care influence the employment of mothers? Findings from western Germany," *Population Research and Policy Review*, 2000, *19* (4), 317–337.

Lefebvre, Pierre and Philip Merrigan, "Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada," *Journal of Labor Economics*, 2008, *26* (3), pp. 519–548.

_ , _ , and Matthieu Verstraete, "Dynamic labour supply effects of childcare subsidies: Evidence from a Canadian natural experiment on low-fee universal child care," *Labour Economics*, 2009, *16* (5), 490–502.

Lengerer, Andrea, Jeanette Bohr, and Andrea Jan"sen, "Haushalte, Familie und Lebensformen im Mikrozensus – Konzepte und Typisierungen," Technical Report, ZUMA-Arbeitsbericht No. 5/2005, Leibniz Institute for the Social Sciences (gesis), Mannheim 2005.

Lotze, Sabine and Holger Breiholz, "Zum neuen Erhebungsdesign des Mikrozensus. Teil 1.," *Wirtschaft und Statistik*, 2002a, *5/2002*, 359–366.

Lundin, Daniela, Eva Mörk, and Björn Öckert, "How far can reduced child-care prices push female labour supply?," *Labour Economics*, August 2008, *15* (4), 647–659.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), pp. 531–542.

Meyers, Marcia K., Theresa Heintze, and DouglasA. Wolf, "Child care subsidies and the employment of welfare recipients," *Demography*, 2002, *39* (1), 165–179.

Michalopoulos, Charles and Philip K. Robins, "Employment and child-care choices in Canada and the United States," *Canadian Journal of Economics*, May 2000, *33* (2), 435–470.

_ and _ , "Employment and child-care choices of single-parent families in Canada and the United States," *Journal of Population Economics*, 2002, *15* (3), 465–493.

_ , _ , and Irwin Garfinkel, "A Structural Model of Labor Supply and Child Care Demand," *Journal of Human Resources*, 1992, *27* (1), 166–203.

Mühler, Grit, "Consequences of the Mixed Provision of Childcare – an Overview on the German Market," Technical Report, ZEW Discussion Paper No. 08-077, Centre for European Economic Research (ZEW) 2010.

Müller, Kai-Uwe and Katharina Wrohlich, "Two Steps Forward – One Step Back? Evaluating Contradicting Child Care Policies in Germany," Technical Report, SOEPpapers on Multidisciplinary Panel Data Research No. 684, German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung, DIW), Berlin 2014.

_ , C. Katharina Spieß, and Katharina Wrohlich, "Kindertagesbetreuung: Wie wird ihre Nutzung beeinflusst und was kann sie für die Entwicklung von Kindern bewirken?," *Vierteljahrshefte zur Wirtschaftsforschung*, 2014, *No. 83*, 49–67.

Nollenberger, Natalia and Núria Rodríguez-Planas, "Child Care, Maternal Employment and Persistence: A Natural Experiment from Spain," Technical Report, IZA Discussion Paper No. 5888, Institute for the Study of Labor (Forschungsinstitut zur Zukunft der Arbeit, IZA), Bonn 2011.

Parera-Nicolau, Antonia and Karen A. Mumford, "Labour Supply and Childcare for British Mothers in Two-Parent Families: A Structural Approach," IZA Discussion Papers 1908, Institute for the Study of Labor (IZA) December 2005.

Powell, Lisa M., "The Impact of Child Care Costs on the Labour Supply of Married Mothers: Evidence from Canada," *Canadian Journal of Economics*, 1997, *30* (3), 577–594.

_ , "Joint Labor Supply and Childcare Choice Decisions of Married Mothers," *Journal of Human Resources*, 2002, *37* (1), 106–128.

Pungello, Elizabeth Puhn and Beth Kurtz-Costes, "Why and How Working Women Choose Child Care: A Review with a Focus on Infancy," *Developmental Review*, 1999, *19* (1), 31–96.

Ribar, David C., "Child Care and the Labor Supply of Married Women: Reduced Form Evidence," *Journal of Human Resources*, 1992, *27* (1), 134–165.

Ribar, David C, "A Structural Model of Child Care and the Labor Supply of Married Women," *Journal of Labor Economics*, July 1995, *13* (3), 558–97.

Schlosser, Amalia, "Public Preschool and the Labor Supply of Arab Mothers: Evidence from a Natural Experiment," Technical Report, The Eitan Berglas School of Economics, Tel Aviv University 2011.

Schober, Pia S. and C. Katharina Spie"s, "Local Day-Care Quality and Maternal Employment: Evidence from East and West Germany," *Journal of Marriage and Family*, 2014, *forthcoming.*

Simonsen, Marianne, "Price of High-quality Daycare and Female Employment," *Scandinavian Journal of Economics*, 09 2010, *112* (3), 570–594.

Spie"s, C. Katharina, "Vereinbarkeit von Familie und Beruf – wie wirksam sind deutsche "care policies"?," *Perspektiven der Wirtschaftspolitik*, 2011, *12*, 4–27.

Tekin, Erdal, "Childcare Subsidies, Wages, and Employment of Single Mothers," *The Journal of Human Resources*, 2007, *42* (2), pp. 453–487.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp, "The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements," *Schmollers Jahrbuch*, 2007, *127* (1), 139–169.

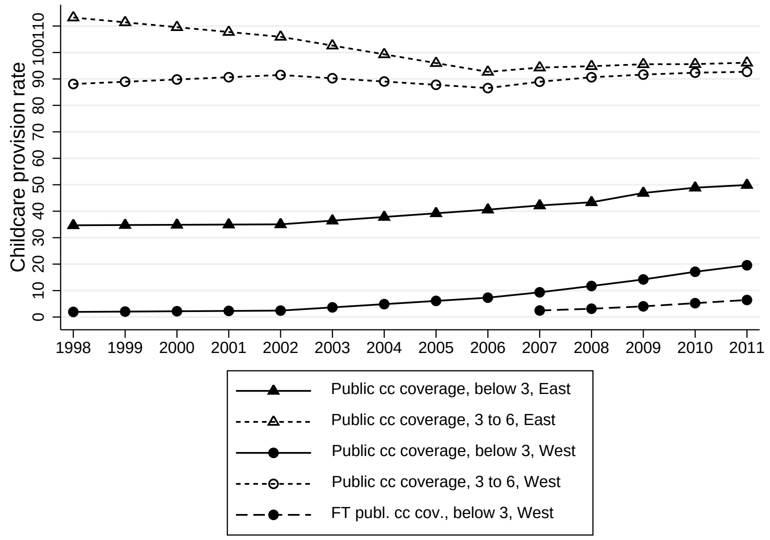
Wrohlich, Katharina, "The excess demand for subsidized child care in Germany," *Applied Economics*, 2008, *40* (10), 1217–1228.

_ , "Labor Supply and Child Care Choices in a Rationed Child Care Market," Technical Report, DIW Discussion Papers No. 1169, German Institute for Economic Research (DIW Berlin), Berlin 2011.

Appendix

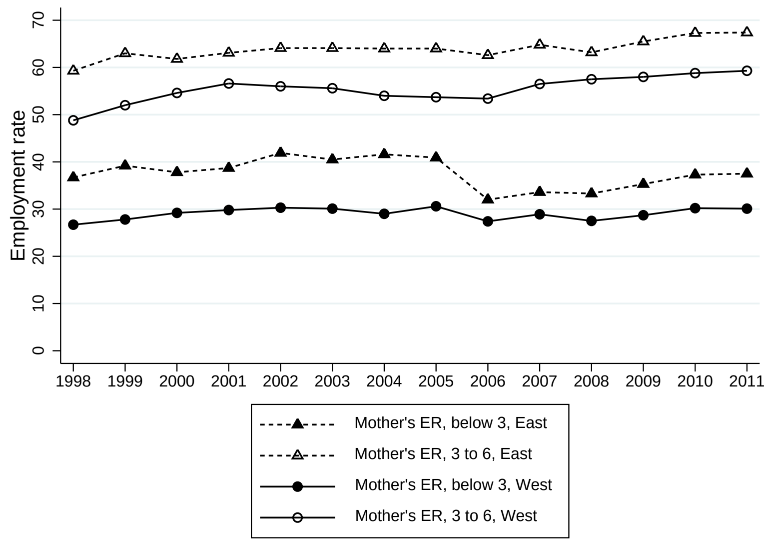
Additional figures

Figure A1: Provision of public childcare in West and East Germany



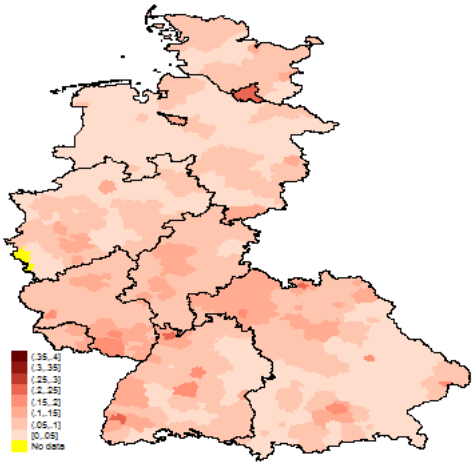
Notes: cc=childcare, below 3=children aged below 3, 3 to 6=children aged 3 to 6, FT=full-time, publ.=public, cov.=coverage, data for full-time care only available from 2002 onwards.
Source: German Statistical Office.

Figure A2: Mothers' employment rates in West and East Germany

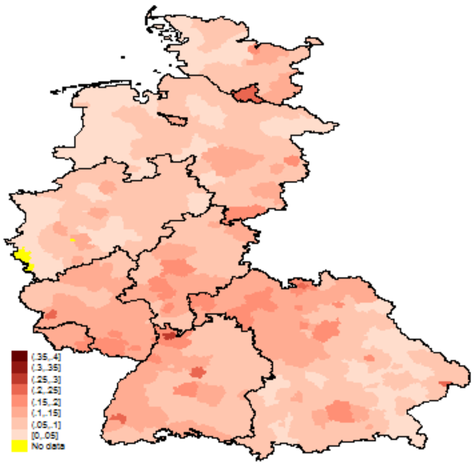


Notes: ER=employment rate, below 3=mothers with children aged 0 to 3, 3 to 6=mothers with children aged 3 to 6.
Source: German Statistical Office, SOEP, MC, own calculations.

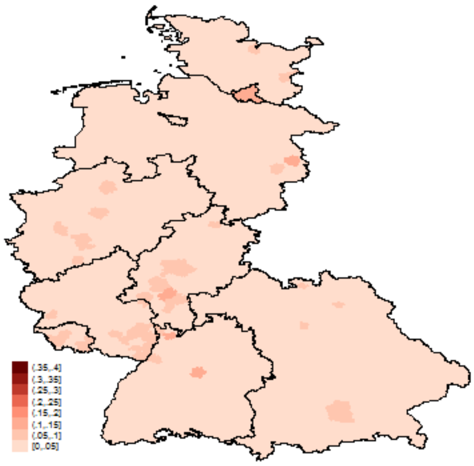
Figure A3: Childcare covarage (in %) at the county level, 2006-2011, West Germany



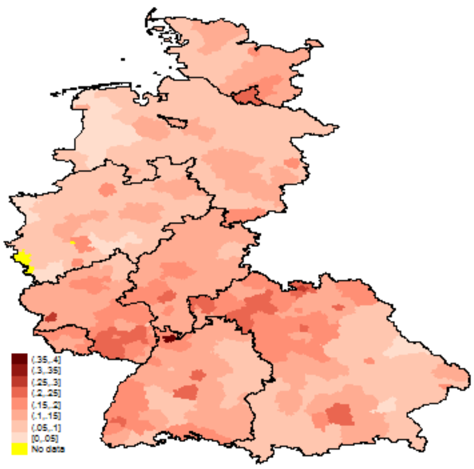
(a) Overall, 2006



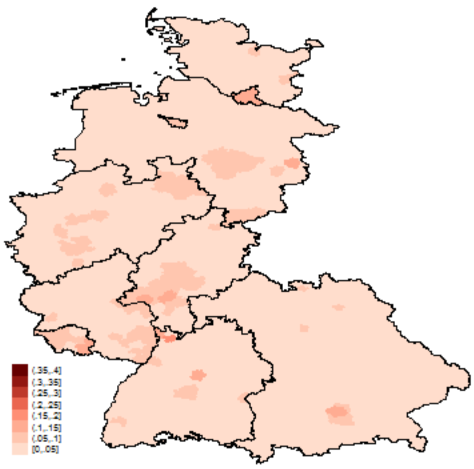
(b) Overall, 2007



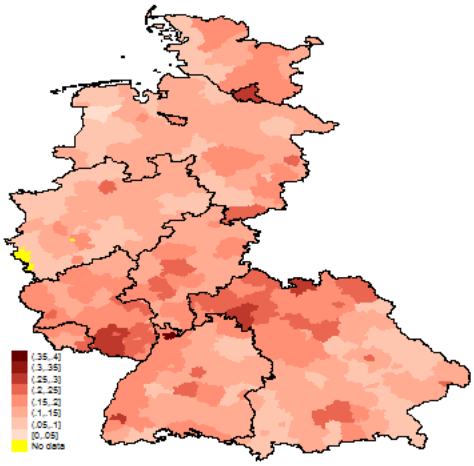
(c) Full-time slots, 2007



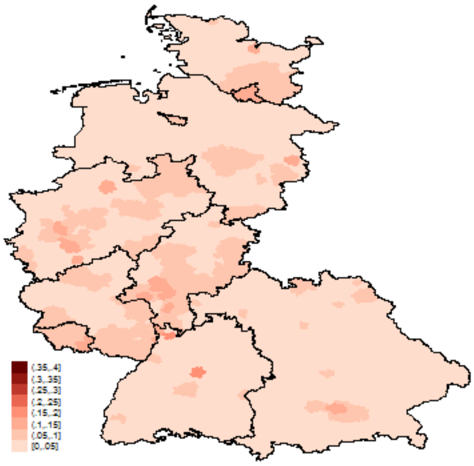
(d) Overall, 2008



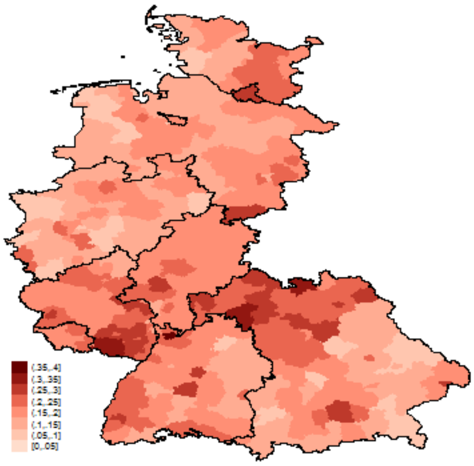
(e) Full-time time slots, 2008



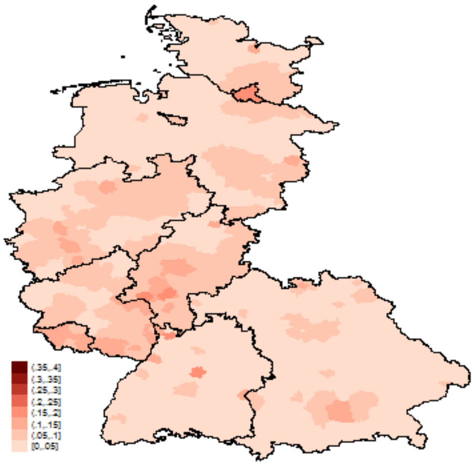
(f) Overall, 2009



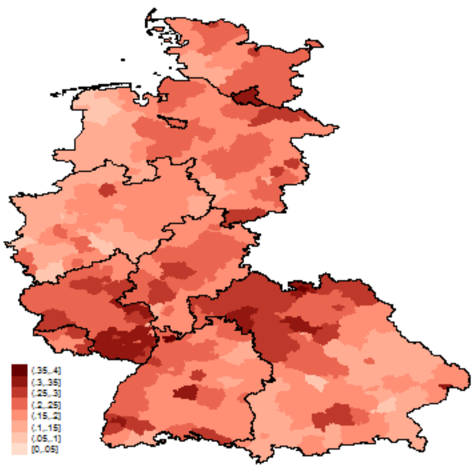
(g) Full-time slots, 2009



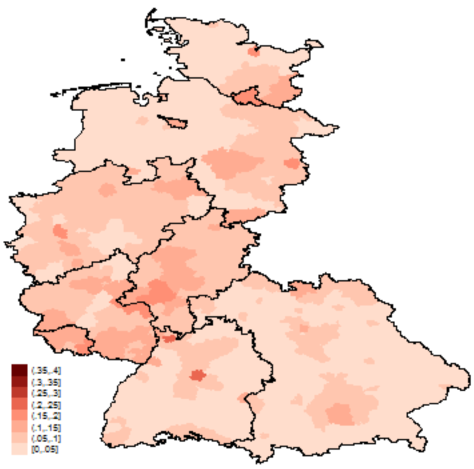
(h) Overall, 2010



(i) Full-time slots, 2010



(j) Overall, 2011



(k) Full-time slots, 2011

Notes: Childcare coverage measured at the county level. Thick lines mark state borders.
Source: German Youth institute, Federal Statistical Office, own calculations.

Additional tables

Table A1: Decriptive statistics, provision of subsidized childcare, 2006-2011

	Overall provision of childcare						Provision of full-time childcare				
	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2007**	**2008**	**2009**	**2010**	**2011**
West Germany											
Mean	0.073	0.094	0.117	0.142	0.171	0.196	0.025	0.032	0.040	0.053	0.064
Minimum	0.010	0.022	0.033	0.037	0.071	0.092	0.000	0.001	0.002	0.004	0.005
Maximum	0.233	0.289	0.352	0.359	0.365	0.376	0.136	0.155	0.166	0.192	0.205
Schleswig-Holst.											
Mean	0.074	0.083	0.114	0.141	0.176	0.212	0.023	0.030	0.040	0.053	0.070
Minimum	0.032	0.038	0.059	0.080	0.084	0.113	0.004	0.005	0.008	0.011	0.019
Maximum	0.141	0.151	0.167	0.192	0.232	0.273	0.071	0.083	0.101	0.134	0.185
Hamburg											
Mean	0.210	0.222	0.229	0.257	0.287	0.324	0.107	0.115	0.136	0.159	0.182
Minimum	0.210	0.222	0.229	0.257	0.287	0.324	0.107	0.115	0.136	0.159	0.182
Maximum	0.210	0.222	0.229	0.257	0.287	0.324	0.107	0.115	0.136	0.159	0.182
Lower Saxony											
Mean	0.048	0.064	0.086	0.114	0.151	0.178	0.014	0.019	0.027	0.037	0.048
Minimum	0.010	0.022	0.035	0.037	0.071	0.092	0.001	0.001	0.003	0.004	0.006
Maximum	0.144	0.160	0.180	0.210	0.264	0.284	0.101	0.113	0.132	0.147	0.184
Bremen											
Mean	0.070	0.083	0.106	0.114	0.138	0.171	0.034	0.044	0.053	0.063	0.092
Minimum	0.036	0.049	0.071	0.079	0.100	0.133	0.022	0.031	0.044	0.050	0.082
Maximum	0.104	0.118	0.141	0.150	0.175	0.209	0.046	0.057	0.061	0.076	0.103
North Rhine-W.											
Mean	0.062	0.064	0.086	0.108	0.133	0.152	0.030	0.040	0.051	0.062	0.072
Minimum	0.028	0.032	0.033	0.062	0.078	0.093	0.006	0.007	0.013	0.025	0.029
Maximum	0.141	0.144	0.184	0.226	0.244	0.250	0.066	0.089	0.116	0.131	0.155
Hesse											
Mean	0.083	0.115	0.135	0.157	0.189	0.209	0.038	0.050	0.064	0.088	0.104
Minimum	0.036	0.074	0.094	0.110	0.137	0.144	0.010	0.012	0.021	0.037	0.051
Maximum	0.141	0.188	0.203	0.216	0.255	0.286	0.108	0.123	0.134	0.178	0.198
Rhineland-Pal.											
Mean	0.096	0.122	0.153	0.178	0.204	0.244	0.034	0.042	0.055	0.072	0.093
Minimum	0.052	0.072	0.090	0.115	0.136	0.147	0.006	0.009	0.016	0.019	0.032
Maximum	0.162	0.202	0.256	0.274	0.330	0.343	0.086	0.093	0.107	0.143	0.179
Baden-Wuertt.											
Mean	0.082	0.110	0.132	0.154	0.179	0.203	0.023	0.029	0.034	0.045	0.054
Minimum	0.023	0.058	0.071	0.087	0.100	0.131	0.001	0.003	0.006	0.009	0.013
Maximum	0.233	0.289	0.352	0.359	0.365	0.376	0.136	0.155	0.166	0.192	0.205
Bavaria											
Mean	0.074	0.098	0.122	0.149	0.179	0.199	0.018	0.022	0.028	0.037	0.045
Minimum	0.018	0.028	0.044	0.061	0.074	0.093	0.000	0.001	0.002	0.004	0.005
Maximum	0.202	0.243	0.272	0.292	0.323	0.357	0.091	0.102	0.109	0.123	0.141
Saarland											
Mean	0.104	0.125	0.149	0.160	0.185	0.210	0.046	0.062	0.083	0.102	0.126
Minimum	0.081	0.103	0.119	0.124	0.150	0.170	0.024	0.036	0.063	0.072	0.097
Maximum	0.139	0.170	0.196	0.204	0.231	0.257	0.085	0.103	0.112	0.144	0.171
Degree of variation – Theil index											
Overall	0.130	0.106	0.079	0.060	0.048	0.042	0.327	0.285	0.247	0.214	0.194
Between states	0.023	0.029	0.020	0.014	0.009	0.009	0.062	0.062	0.062	0.059	0.055
Within states	0.106	0.077	0.059	0.047	0.039	0.032	0.264	0.223	0.185	0.155	0.139

Notes: The Theil index is decomposable into a weighted sum of between- and within-subgroup inequality. For a definition and the relation to ohter inequality measures, see Cowell (2000).
Source: German Youth institute, Federal Statistical Office, own calculations.