Quantifying Okun's Leaky Bucket: The Case of Progressive Childcare Subsidies*

David Koll Dominik Sachs Fabian Stürmer-Heiber Hélène Turon

August 11, 2026

Abstract

We formalise and estimate the dynamic marginal efficiency cost of redistribution (MECR) in the spirit of Okun's "leaky bucket" to compare income-contingent childcare subsidies with the broader tax-transfer system. Because the MECR depends on heterogeneous behavioral responses and fiscal effects, we estimate a structural model of household childcare demand and maternal labor supply for Germany, allowing for informal childcare and leisure. Our analysis identifies two competing forces: labor supply responses increase the MECR of childcare subsidies relative to the tax-transfer system, while child development effects decrease it. Under empirically plausible assumptions, progressive childcare subsidies are the more efficient redistribution tool.

JEL Code: H23, H31, J13, J22, J24.

Keywords: Female labor Supply, Childcare, Family Policies, Fiscal Externalities, Dynamic Discrete Choice, Redistribution

---

*Koll (corresponding author): International Monetary Fund, Washington D.C. (dkoll@imf.org) Sachs: University of St. Gallen and CESifo (dominik.sachs@unisg.ch) Stürmer-Heiber: Allianz. Turon: University of Bristol, CEPR & IZA (helene.turon-lacarrieu@bristol.ac.uk). We thank Árpád Ábahám, Spencer Bastani, Friedrich Breyer, Elizabeth Caucutt, Mark Colas, Mariacristina De Nardi, Maya Eden, Philipp Eisenhauer, Eric French, Hans-Martin von Gaudecker, Nezih Guner, Peter Haan, Anne Hannusch, Emanuel Hansen, Albert Jan Hummel, Bas Jacobs, Mike Keane, Hamish Low, Linus Mattauch, Joseph Mullins, Nicola Pavoni, Maria Polyakova, Daniel Reck, Víctor Ríos-Rull, Karl Schulz, Michèle Tertilt, Petra Todd, Gianluca Violante, Johanna Wallenius, and Hanna Wang for comments and suggestions. We thank Lea Fricke, Tim Hug, Nuan Stahl, and Victoria Szabo for excellent research assistance. We thank participants at the CEPR Public Economics, CESifo Public Economics Week & Conference, SED, Barcelona Summer Forum, ZEW Public Finance, FROGEE, MadMac, VfS, NASMES, IIPF, DSE Conference, CRC 190 & 224 Retreats as well as at HEC Lausanne, UC3 Madrid, CRED Paris, UAB, Bank of Canada, IAB, Copenhagen, Tinbergen for helpful comments. David Koll gratefully acknowledges funding by the German Research Foundation (CRC TR 224 & Leibniz-Prize) and Fabian Stürmer-Heiber through the Doctoral Program "Evidence-Based Economics". This paper was previously circulated as "Equity and Efficiency of Childcare Subsidies: A Dynamic Structural Approach". **Disclaimer:** The views expressed in this paper are solely those of the authors and do not necessarily represent the views of the International Monetary Fund, its Executive Board, or its management.

1 Introduction

> *"However, the program has an unsolved technological problem: the money must be carried from the rich to the poor in a leaky bucket. Some of it will simply disappear in transit, so the poor will not receive all the money that is taken from the rich... Of course the leak presents an inefficiency. The inefficiencies of real world redistribution include the adverse effects on the economic incentives of the rich and the poor... " Okun (1975, p. 89)*

Public childcare is a major in-kind family policy in advanced economies, and in a broad range of countries the support that families receive is explicitly directed towards lower-income households through income-related subsidies (European Commission/EACEA/Eurydice 2019, OECD 2020). This paper studies German childcare subsidies and asks how efficiently they transfer resources from higher-income to lower-income families relative to the broader tax-transfer system. Public childcare spending in Germany amounted to €37bn in 2019 – slightly more than 1% of GDP. This spending is redistributive: parents at the 80th percentile of the household income distribution pay about 60% more for full-time public childcare for a child under 3 years old than families at the 20th percentile.[1]

By making childcare fees depend on household income, governments embed a redistributive component in childcare policy. This redistributive margin operates alongside a broader tax-transfer system that already redistributes substantially and includes child-specific elements such as child benefits and child tax allowances. Governments thus use multiple instruments to redistribute resources from higher- to lower-income families with young children. A natural question is therefore whether the current progressivity of the childcare fee schedule is a relatively cheap or a relatively costly way of doing so, compared with the alternative redistributive margins already present in the tax-transfer system. Put differently: is it more efficient to give lower-income families cheaper childcare, or to give them money through the tax-transfer system? We do not ask whether redistribution should be expanded or retrenched overall, nor whether childcare should replace the tax-transfer system as the main vehicle for redistribution. Our question is more targeted: at the *current level of redistribution* embedded in each instrument, which one redistributes more efficiently at the margin?

We tackle this question by introducing the marginal efficiency cost of redistribution (MECR), a metric that operationalizes the leaky-bucket idea of Okun (1975). For any given policy instrument, the MECR quantifies the leakage associated with redistributing the marginal euro: how much reaches lower-income families and how much is lost through behavioral and fiscal responses. To quantify the MECR, we consider small, dynamically budget-neutral perturbations that marginally steepen or flatten existing schedules and use the resulting leakage to assess efficiency of the redistributive margins already present in the status quo. These perturbations are therefore not reform proposals in their own

[1] Authors' calculations.

right, but an evaluation device for the redistribution built into real-world schedules as they currently operate. This makes the MECR especially useful in policy environments where governments redistribute through several instruments at once. In the present application, we compare redistribution through income-contingent childcare fees with redistribution through the tax-transfer system for families with young children. The comparison is informative whether policymakers wish to expand or to retrench redistribution across families: redistribution should be expanded through the instrument with the lower MECR and scaled back through the one with the higher MECR.

Two dynamic channels. Any application of the MECR requires a judgement about which economic responses are first-order for the policy under study. In the case of childcare, two dynamic channels stand out. The first is maternal labor supply, together with the career effects that arise because shorter labor-market interruptions translate into higher future earnings and future tax payments. The second is child development: by changing children's exposure to formal care, childcare policy may affect later-life outcomes and thus the long-run consequences of redistribution (Busse and Gathmann 2020, Cornelissen et al. 2018, Havnes and Mogstad 2011a). Our objective is to incorporate both channels into the MECR analysis.

Empirical strategy. A central challenge is that no single dataset allows these margins to be identified jointly within one unified framework. We therefore adopt a modular empirical strategy that we now describe.

For maternal labor supply and dynamic career effects, a dynamic structural model is well suited, and German panel survey data allow for careful estimation. The MECR depends on behavioral responses *across the income distribution* — specifically, on heterogeneous own- and cross-price elasticities of labor supply and childcare demand. Simple aggregate elasticities would be a poor substitute: they average the responses of rich and poor households, which enter the MECR formula in different ways. The reduced-form evidence needed to recover exactly these heterogeneous objects is not readily available. A dynamic model also allows us to incorporate career effects: lower maternal labor supply today may reduce future wages and future tax payments, and these long-run fiscal effects turn out to be quantitatively important. We therefore estimate a dynamic structural model of household labor supply and childcare demand in which families differ in wages, family composition, preferences for maternal care and leisure, and access to informal childcare. We cross-validate the quantified model against existing reduced-form evidence on labor supply elasticities and childcare demand responses (Busse and Gathmann 2020, Gathmann and Sass 2018).

For children's long-run outcomes, the empirical challenge is different and we are deliberately agnostic. Rather than embed a structural child-development block inside the model — which would require additional assumptions that are difficult to discipline with the

available data — we proceed in two steps. We first use the structural model to predict how changes in the fee schedule affect childcare attendance, and then map these policy-induced changes into later earnings using a range of possible values for the returns to childcare. Critically, this evidence is scarce and context-specific. While German studies document heterogeneous short-run child-development effects of childcare (Busse and Gathmann 2020, Cornelissen et al. 2018), the long-run earnings returns in a German institutional setting with broad access to publicly regulated provision are not reliably known. We therefore report results across a full range of assumptions on the returns to childcare, and – crucially – we show transparently how large these returns would need to be for the ranking of redistributive instruments to reverse. We view this transparency as valuable, because it shows clearly where additional evidence would matter for policy conclusions. In this respect, our approach is close in spirit to a broader modern policy-evaluation tradition, including the marginal value of public funds (MVPF) literature: different links in the causal chain are disciplined by the sources of evidence that are most informative for them (Hendren and Sprung-Keyser 2020).

Quantitative results. Our benchmark results, which ignore child development effects, show that the MECR of the progressive childcare subsidy schedule is 0.42, while that of the tax-transfer schedule is lower at 0.28. For each euro taken from above-median income families, 58 cents (respectively 72 cents) reach below-median families. The higher leakage from childcare subsidies is largely driven by cross-price effects: making childcare more expensive for high-wage mothers and cheaper for low-wage mothers generates negative labor supply responses among the former that are fiscally larger than the positive responses among the latter. Dynamic wage effects matter not only for the levels of the two MECRs, but also for their difference: they amplify the fiscal impacts of both instruments by around 40% and explain just over 20% of the gap between them.

When we incorporate child development, the ranking can reverse. Under the empirically plausible assumption that children from lower-income households benefit more from childcare attendance than those from higher-income households, the MECR of childcare subsidies falls significantly. For returns to one year of childcare attendance of 3% for below-median children and 0% for above-median children, the MECRs of the tax-transfer schedule and childcare subsidies become 0.31 and 0.24 respectively, making childcare subsidies the more efficient redistributive tool. A return of 2.02% for below-median children (with 0% for above-median) is sufficient to equalise the two instruments. We also show that our methodology extends straightforwardly to other redistribution thresholds beyond the median.

Contribution to the literature. This paper connects structural work on household decision-making with the public finance literature on optimal redistribution. On the optimal tax side, Domeij and Klein (2013) establish in a Ramsey setting that childcare

should be subsidized to counteract negative incentive effects of labor taxes, finding a linear subsidy of around 50% optimal for Germany. Bastani, Blomquist, and Micheletto (2020) show in a Mirrleesian setting that heterogeneous childcare quality weakens the subsidisation argument, while Ho and Pavoni (2020) characterise optimal income-varying childcare subsidies for the US, finding the optimal schedule should decrease more steeply with income than current policy. Our paper differs from this literature in a fundamental respect: rather than characterising an optimal reform, we take a more applied perspective and evaluate the efficiency of redistribution embedded in existing instruments at current policy settings. This is directly informative about which existing instrument provides the cheaper margin along which to adjust redistribution. Our local perturbations keep the comparison tied to the current policy environment, while still using the structural model to account for the behavioral margins that matter for redistribution.

Our paper also connects with a literature that estimates dynamic labor supply responses to family policy (Guner, Kaygusuz, and Ventura 2020; Blundell et al. 2016; Attanasio, Low, and Sánchez-Marcos 2008 ; Bick 2016; Wang 2022). A key common feature is the depreciation of human capital during non-participation and lower returns to part-time versus full-time experience — dynamic wage effects that prove quantitatively important in our MECR calculations. We build on and extend Turon (2019)'s dynamic discrete-choice model of couples with children, adding leisure, informal childcare access, and a richer heterogeneity structure. Related to us, Mullins (2022) studies the optimal design of cash transfers to single mothers in a structural household model that also incorporates social mobility through endogenous future earnings of children.

Our paper makes two core contributions relative to this literature. First, we introduce the MECR as a practical tool for evaluating the redistributive efficiency of policies *already in place*, providing policymakers with a clear metric to identify which existing instruments redistribute more cheaply at the margin. Second, by applying the concept to both childcare subsidies and the tax-transfer system jointly, we offer what amounts to a direct answer to the question of whether it is more efficient to give lower-income families cheaper childcare or more money — while being explicit about the conditions under which each answer holds.

The structure of the paper is as follows: Section 2 introduces the methodological concept of the MECR in a simple static model. We outline the dynamic model in Section 3. Section 4 discusses the quantification of the model. We describe the model fit in Section 5 and the quantitative results in Section 6. Section 7 concludes.

2 Formalising the MECR in a simple model

In this section, we introduce a static model to clarify the core principles that determine the marginal efficiency cost of redistribution for childcare subsidies and the tax-transfer

system. We formally define our measure of the MECR and highlight the underlying trade-offs theoretically. This analysis uncovers the complexity of the factors contributing to the MECR, emphasising the need for the dynamic structural model used in subsequent sections.

2.1 Parents' preferences, constraints and decisions

Choices and constraints. For this simple model, we examine a heterogeneous group of households, each with two parents and one young child. There is one unit of time. We assume that male labor supply is fixed at one unit, i.e. full time. For women, decisions are endogenous, requiring them to allocate their single unit of time across domestic childcare D, labor supply H, or leisure L. Formally, mothers face a time budget constraint

$$L + H + D = 1. \tag{1}$$

In Section 2, we assume that children require care for the full unit of time and that there is no private childcare available. Hence, children's time is shared between domestic childcare, D and time spent with a public childcare provider (nursery), N:

$$D + N = 1. \tag{2}$$

Preferences and heterogeneity. Denote each household by i. We assume that utility is quasi-linear in consumption C and heterogenous across households i: $C + u_i(L, D)$.[2] Households value the mother's leisure and the amount of childcare that she provides, for reasons detailed above. The quasi linearity implies that childcare and labor supply decisions only depend on relative prices but not on wealth. We make this assumption here for simplicity to better single out the most important forces, as is often done in the optimal tax literature. In our dynamic structural model, however, we will consider a utility function with income effects.

Besides preferences, households differ in their wages $(w_{f,i}, w_{m,i})$, where subscripts f and m denote female and male. Household income is defined as $y_i = w_{f,i}H_i + w_{m,i}$.

Public policies. Public childcare is available at price $\mathcal{K} - s(y_i)$ per unit. $\mathcal{K}$ represent the cost per unit of childcare and $s(y_i)$ is the subsidy per unit of childcare, which depends on household income y_i. Further, households face a nonlinear tax-transfer system $T(y_i)$. $T(y_i)$ can be negative which reflects that the household is a transfer recipient. Note that this also incorporates child-dependent policies that do not depend on the mode of childcare,

[2]The term $u_i(L, D)$ may capture the parents' utility derived from leisure L and time with their children D and an 'altruistic' term which, e.g., captures the impact that the form of care has on the later earnings of their children and therefore ultimately utility of their children. We do not make any a priori assumptions about altruism and also allow it to be heterogeneous across families.

e.g., the child component of a welfare benefit or child tax credits. Since all households have exactly one child in this first model, this child-dependence does not show up in $T(y_i)$.

Decision problem. Household i solves the following problem:

$$\max_{H_i,N_i} C_i + u_i\left(L_i, D_i\right) \quad \text{subject to} \quad \begin{cases} C_i = y_i - T(y_i) - \left[\mathcal{K} - s(y_i)\right] N_i, \\ y_i = w_{f,i} H_i + w_{m,i} \text{ , (1) and (2).} \end{cases}$$

The first-order condition for H_i is given by:

$$\left[1 - T'(y_i) + s'(y_i) N_i\right] w_{f,i} = \frac{\partial u_i}{\partial L_i}.$$

This shows the trade-off between work and leisure, holding D constant. We see that labor supply is distorted by both the marginal tax rate T' and the marginal childcare subsidy s'. We define the "labor wedge" τ_i^H and the implied net wage $w_{f,i}^{net}$ as:

$$\tau_i^H = T'(y_i) - s'(y_i) N_i \quad \text{and} \quad w_{f,i}^{net} = (1 - \tau_i^H) w_{f,i}. \tag{3}$$

Next, we turn to the first-order condition for N:

$$\frac{\partial u_i}{\partial D_i} + \mathcal{K}\left[1 - \frac{s(y_i)}{\mathcal{K}}\right] = \frac{\partial u_i}{\partial L_i}.$$

This illustrates the trade-off between leisure L and domestic childcare D holding labor supply H constant. The left-hand side shows the utility gains from consuming what one would have spent on one hour of public childcare, $\mathcal{K}\left[1 - \frac{s(y_i)}{\mathcal{K}}\right]$, and the marginal utility from domestic childcare $\frac{\partial u_i}{\partial D}$. The utility costs in terms of foregone leisure are on the right-hand side. Similar to the labor wedge τ_i^H, we now define the "childcare wedge" τ_i^N, which is the rate of subsidisation for one hour of public childcare, and the implied net childcare cost $\mathcal{K}_i^{net}$:

$$\tau_i^N = \frac{s(y_i)}{\mathcal{K}} \quad \text{and} \quad \mathcal{K}_i^{net} = \left(1 - \tau_i^N\right) \mathcal{K}. \tag{4}$$

The net-prices $w_{f,i}^{net}$ and $\mathcal{K}_i^{net}$ affect both the labor supply and the childcare demand decisions. (3) and (4) show that the subsidy schedule affects both of these net-prices: the marginal subsidy $s'(y)$ affects the net wage (3) and the absolute subsidy $s(y)$ affects the net childcare cost (4). This contrasts with the second policy instrument, the tax-transfer schedule, which only affects the net-wage through the marginal tax rate $T'(y)$. This asymmetry is a key reason for the difference in efficiency costs of redistribution between the two instruments, as we show below.

2.2 Measuring the marginal efficiency cost of redistribution

To quantify the marginal efficiency cost of redistribution, we introduce parametric hypothetical perturbations for both policy instruments. The changes are budget neutral by construction and redistribute from households above a given percentile p of the income distribution to households below that percentile. To understand our notion of budget neutrality first note that we define net revenue NR, i.e., tax revenue net of subsidy spending, as:

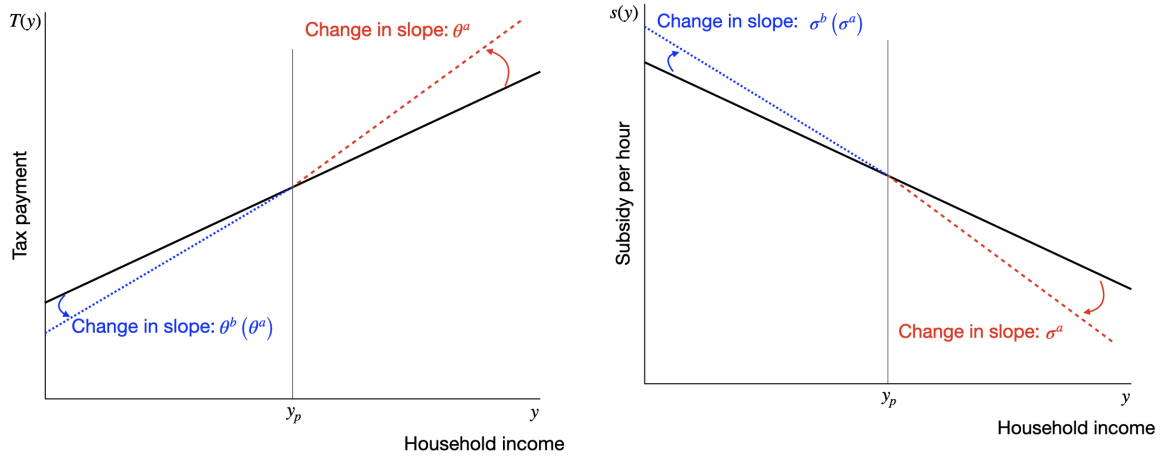
$$NR = \int_i T(y_i)di - \int_i N_i s(y_i)di. \tag{5}$$

We then define budget neutral perturbations which imply that $d(NR) = 0$ after accounting for all behavioral changes. In the static model of this section, budget neutrality refers to contemporaneous net revenue. In the quantitative dynamic model below, we impose budget neutrality in net-present-value terms. These simple perturbations allow us to derive how costly the last unit of redistribution of a given real world policies is, in terms of efficiency costs.

Perturbation of the tax-transfer schedule. Before formally defining the perturbation, we first provide an illustration in Figure 1a. The black solid line represents the initial tax-transfer schedule, i.e. tax payment net of transfers as a function of household income. Note that we depict the initial tax-transfer schedule as linear merely for simplicity – our analysis does not necessitate a linear baseline tax-transfer schedule. The change increases the marginal tax rate and the absolute tax payment for households with incomes above y_p, as shown by the red dashed line. For households with incomes below y_p, the marginal tax rate is also increased, but the absolute tax payment decreases, as indicated by the blue dotted line. Given a specific increase in tax payments above the cutoff income y_p (a given red dashed line), the blue dotted line is chosen so that, after accounting for all household responses, the entire perturbation remains budget-neutral. This change redistributes resources, i.e. the marginal euro, from households with incomes above y_p to those below y_p. It can be formally defined as follows:

$$\widehat{T}_p(y) = \begin{cases} \theta^a \left[y - y_p\right] & \text{for } y > y_p \\ -\theta^b\left(\theta^a\right)\left[y_p - y\right] & \text{for } y \leq y_p \end{cases} \tag{6}$$

θ^a is the increase in the marginal tax rate above y_p and $\theta^b\left(\theta^a\right)$ is the increase in the marginal tax rate below y_p. We consider small perturbations with $\theta^a \to 0$ so that we can focus on first-order effects in our analysis. As indicated above, for a given value of θ^a, $\theta^b\left(\theta^a\right)$ is defined such that the change is budget neutral ($dNR = 0$), once agents have adapted their behavior. The perturbed tax-transfer schedule is then given by $T(y) + \widehat{T}_p(y)$.



(a) Tax-transfer perturbation

(b) Childcare subsidy perturbation

Figure 1: Illustrations of tax and subsidy changes

Perturbation of the childcare subsidy schedule. The perturbation of the childcare subsidy schedule, which we illustrate in Figure 1b, is defined analogously:[3]

$$\hat{s}_p(y) = \begin{cases} -\sigma^a (y - y_p) & \text{for } y > y_p \\ \sigma^b(\sigma^a) (y_p - y) & \text{for } y \leq y_p \end{cases} \tag{7}$$

where superscripts a and b denote above and below y_p. The perturbed hourly subsidy schedule is given by $s(y)+\hat{s}_p(y)$. This change implies that the subsidy per hour is decreased above income level y_p and increased below. Marginal subsidies – i.e. how the subsidy per hour decreases with household income – are increased (in absolute terms) throughout the income distributions. The subsidy perturbation (7) is very similar to the tax-transfer change (6) in its distributional effects; the difference arises because the financial impact of the subsidy change on household i depends on not only on y_i but also on N_i. However, as noted above already, the incentive effects of the perturbation are different: while the tax-transfer change only implies changes in the net wage, the subsidy change changes both the net wage and the net childcare cost.

Relation to elementary tax reforms. Following the Mirrleesian approach of optimal nonlinear taxation, a so-called "elementary tax reform" could be considered the natural reform for our purpose.[4] In elementary tax reforms, the marginal tax rate (or the marginal rate of subsidisation) is altered only within a very narrow interval surrounding some income level y_p. The additional tax revenue generated can then be redistributed in a lump-sum fashion, which implies – similar to our perturbations – redistribution from those above y_p to those below y_p.

[3]The functions $\theta^b(\cdot)$ and $\sigma^b(\cdot)$ also depend on the percentile p, which we omit for ease of notation.

[4]See Piketty (1997), Saez (2001), Golosov, Tsyvinski, and Werquin (2014) and Sachs, Tsyvinski, and Werquin (2020).

As put forward, for example, by Saez and Stantcheva (2018), the marginal efficiency costs of redistribution implied by these reforms are highly sensitive with respect to the value of the local Pareto parameter of the continuous income distribution. This sensitivity arises because marginal tax rates are essentially increased at a single point in the income distribution. This makes results somewhat sensitive to the coarseness of the grid, see Heathcote and Tsujiyama (2021) for a detailed elaboration. Our proposed perturbations $\widehat{T}_p(y)$ and $\hat{s}_p(y)$ circumvent this complication and results are hence less sensitive w.r.t. the coarseness of the grid. The MECR is a versatile tool that can also be applied to elementary tax reforms, see e.g., Fricke and Sachs (2024).

2.3 Marginal efficiency cost of redistribution: definition

We now formalise the popular idea of Okun's leaky bucket based on our parametric perturbations (6) and (7). Our definition is based on money-metric utility changes, i.e. compensating variations (Hicks 1939), often labeled as willingness to pay (Hendren and Sprung-Keyser 2020). These are not equal to changes in consumption: they are the mechanical changes in consumption that would arise for fixed behavior and have first-order effects on individual utility. Changes in income and consumption induced by changes in choices have only second-order effects on utility due to the envelope theorem.

Definition 1. *The marginal efficiency cost of redistribution of a tax-transfer change $\widehat{T}_p$ as defined in* (6) *or a subsidy change $\hat{s}_p$ as defined in* (7) *is given by:*

$$MECR(\widehat{T}_p) = 1 - \frac{CV^b\left(\widehat{T}_p\right)}{\left|CV^a\left(\widehat{T}_p\right)\right|} \tag{8}$$

$$MECR(\hat{s}_p) = 1 - \frac{CV^b\left(\hat{s}_p\right)}{\left|CV^a\left(\hat{s}_p\right)\right|} \tag{9}$$

where

$$CV^a(\widehat{T}_p) = -\theta^a \int_{i:y_i>y_p} [y_i - y_p]\, di \quad \textit{and} \quad CV^b(\widehat{T}_p) = \theta^b(\theta^a) \int_{i:y_i\leq y_p} [y_p - y_i]\, di.$$

are the aggregated compensating variations of households with income above and below y_p for the tax-transfer perturbation $\widehat{T}_p$. Further,

$$CV^a(\hat{s}_p) = -\sigma^a \int_{i:y_i>y_p} N_i\,[y_i - y_p]\, di. \quad \textit{and} \quad CV^b(\hat{s}_p) = \sigma^b(\sigma^a) \int_{i:y_i\leq y_p} N_i\,[y_p - y_i]\, di.$$

are the aggregated compensating variations for the subsidy perturbation $\hat{s}_p$.

The marginal efficiency cost of redistribution as defined in (8) and (9) can be interpreted as follows: for each unit of aggregate money-metric utility taken from the group earning

above y_p, $(1 - MECR)$ units of aggregate money-metric utility gain can be achieved for the group earning below y_p and $MECR$ is lost through the leakage of Okun's bucket.

If households did not change their behavior as a response to the perturbations, we would have $MECR(\widehat{T}_p) = MECR(\hat{s}_p) = 0$. In other words, each mechanical Euro taken from those above the p-th percentile would reach those below p. If however households respond to the changes, the $MECR$ are no longer zero because behavioral changes affect net revenue (5). These changes in net revenue due to changes in behavior are referred to as *fiscal externalities* in the public finance literature (Hendren 2016) and determine the $MECR$.

To obtain closed-form expressions for $MECR(\widehat{T}_p)$ and $MECR(\hat{s}_p)$ in terms of (10)-(15), we need to solve for the budget-neutral values of $\theta^b(\theta^a)$ and $\sigma^b(\sigma^a)$. We derive such analytical expressions for $MECR(\widehat{T}_p)$ and $MECR(\hat{s}_p)$ in Online Appendix A. Before turning to the most important aspects of these results in Section 2.4, we briefly relate our concept to the marginal value of public funds.

Relation to marginal value of public funds. Our concept of the MECR has a direct relation to the recent and widely applied concept of the marginal value of public funds (MVPF) introduced by Hendren and Sprung-Keyser (2020). The MVPF measures the aggregate compensating variation of a policy divided by the net cost of the policy, where the measure of costs accounts for the fiscal externalities of the policy. To understand the relation between MECR and MVPF, it is useful to split our perturbation into two parts. E.g., for the tax-transfer change, denote the tax increase above y_p by $\widehat{T}_p^a$ and the tax decrease below y_p by $\widehat{T}_p^b$. Denote the respective MPVFs as $MVPF(\widehat{T}_p^a)$ and $MVPF(\widehat{T}_p^b)$. It is then simple to show that:[5]

$$MECR(\widehat{T}_p) = 1 - \frac{MVPF(\widehat{T}_p^b)}{MVPF(\widehat{T}_p^a)}.$$

We consider our MECR as complementary to the MVPF. It allows to compare different redistribution tools regarding their efficiency in achieving redistribution. Below we provide a showcase for the implementation in a structural model. Our MECR measure, similarly to the MVPF, can be quantified based on given empirical estimates if they are available.

2.4 Marginal efficiency cost of redistribution: theoretical results

In a first step, we introduce four terms that capture the fiscal externalities of changes in net wage and net cost.

[5]Since our perturbation $\widehat{T}_p$ is budget neutral, the net-fiscal effects in the denominators of $MVPF(\widehat{T}_p^b)$ and $MVPF(\widehat{T}_p^a)$ cancel out. Hence, what remains of the ratio $\frac{MVPF(\widehat{T}_p^b)}{MVPF(\widehat{T}_p^a)}$ are the compensating variations, which then gives (8).

Definition 2. *An increase in the net wage $w_{f,i}^{net}$ implies a fiscal externality through an own-price effects O_w^i on hours worked and a cross-price effect X_w^i on childcare demand:*

$$O_w^i = \frac{\tau_i^H}{1-\tau_i^H} \cdot y_{f,i} \cdot \varepsilon_{H,w_f^{net}}^i > 0 \tag{10}$$

$$X_w^i = -\frac{\tau_i^N}{1-\tau_i^H} \cdot N_i \cdot \varepsilon_{N,w_f^{net}}^i < 0 \tag{11}$$

An increase in the net cost $\mathcal{K}_i^{net}$ implies a fiscal externality through an own-price effect $O_{\mathcal{K}}^i$ on childcare demand and a cross-price effect $X_{\mathcal{K}}^i$ on hours worked:

$$O_{\mathcal{K}}^i = \frac{1}{\mathcal{K}} \cdot \frac{\tau_i^N}{1-\tau_i^N} \cdot N_i \cdot |\varepsilon_{N,\mathcal{K}^{net}}^i| > 0 \tag{12}$$

$$X_{\mathcal{K}}^i = -\frac{1}{\mathcal{K}} \cdot \frac{\tau_i^H}{1-\tau_i^N} \cdot y_{f,i} \cdot |\varepsilon_{H,\mathcal{K}^{net}}^i| < 0 \tag{13}$$

where $\varepsilon_{Z,I}^i = \frac{\partial log(Z_i)}{\partial log(I_i)}$ for choices $Z_i = H_i, N_i$ and instruments $I_i = w_{f,i}^{net}, \mathcal{K}_i^{net}$.

The derivation of these four objects can be found in Online Appendix A. Note that all four fiscal externalities are proportional to the product of the relevant elasticity and the baseline value of N_i and $y_{f,i}$ respectively.

Since a higher net wage implies higher labor supply and therefore higher tax revenue, the own-price wage effect, O_w^i is positive. On the other hand, the cross-price wage effect, X_w^i is negative since an increase in the net wage increases childcare demand, which itself increases government spending on subsidies.

The fiscal externality coming from the own-price childcare effect, $O_{\mathcal{K}}^i$ is positive since the increase in net cost implies a reduction in childcare demand, which reduces subsidy spending. Finally, the cross-price effect of the net cost change, $X_{\mathcal{K}}^i$, is negative because a higher net cost of childcare implies lower labor supply and therefore lower tax revenue. Since X_w^i and $X_{\mathcal{K}}^i$ are negative, the propositions below state comparative statics for these cross-price effects in terms of their absolute values, $|X_w^i|$ and $|X_{\mathcal{K}}^i|$. First, we summarise how the tax-transfer perturbation affects the net wage (3) and the net childcare cost (4).

Lemma 1. *A tax-transfer perturbation $\widehat{T}_p(y)$ as defined in (6)*

1. *lowers the net wage $w_{f,i}^{net}$ for all households.*
2. *does not affect the net childcare costs $\mathcal{K}_i^{net}$ for any household.*

Hence, to calculate the $MECR$ of the tax-transfer change, we only need to take into account that households adjust their behavior due to a change in their net wage $w_{f,i}^{net}$.

Proposition 1. *The marginal efficiency cost of redistribution of redistribution $MECR(\widehat{T}_p)$ is*

1(a) increasing in $E[O_w|y > y_p]$ *and* $E[O_w|y < y_p]$.

1(b) decreasing in $E[|X_w| \mid y > y_p]$ *and* $E[|X_w| \mid y < y_p]$.

Proof. This immediately follows from the formula for $MECR(\widehat{T}_p)$ in (26) in Online Appendix A. □

1(a) refers to the own-price effect of a change in w_f^{net}: all households work less due to the lower net-wage. This results in a negative fiscal externality, which increases the MECR of the tax-transfer change. 1(b) refers to the magnitude of the cross-price effect of this change in w_f^{net}: households will also demand less childcare, which results in a positive fiscal externality leading to lower MECR.

We now turn to a childcare subsidy perturbation.

Lemma 2. *A subsidy perturbation* $\hat{s}_p(y)$ *as defined in (7):*

1. *lowers the net wage* $w_{f,i}^{net}$ *for all households.*
2. *increases (resp. decreases) the net childcare cost* $\mathcal{K}_i^{net}$ *for households with income above (resp. below)* y_p.

As for the tax-transfer change, a steeper slope of the childcare subsidy schedule lowers the net wage, this is captured by point 1 in Lemma 2. The second point captures the change in the net childcare cost incurred by the perturbation above and below percentile p. Note that the change in the price per hour is larger the further away the household's income is from y_p. The following proposition states how these changes in net wages and net childcare costs affect $MECR(\hat{s}_p)$:

Proposition 2. *The marginal efficiency cost of redistribution* $MECR(\hat{s}_p)$ *is:*

1(a) increasing in $E[O_w \cdot N|y > y_p]$ *and* $E[O_w \cdot N|y < y_p]$.

1(b) decreasing in $E[|X_w| \cdot N \mid y > y_p]$ *and* $E[|X_w| \cdot N \mid y < y_p]$.

2(a) decreasing in $E[O_{\mathcal{K}} \cdot (y_i - y_p)|y > y_p]$ *and increasing in* $E[O_{\mathcal{K}} \cdot (y_p - y_i)|y < y_p]$.

2(b) increasing in $E[|X_{\mathcal{K}}| \cdot (y_i - y_p) \mid y > y_p]$ *and decreasing in* $E[|X_{\mathcal{K}}| \cdot (y_p - y_i) \mid y < y_p]$.

1(a) and 1(b) mirror the statements in Proposition 1 and capture the own-price effect and the magnitude of the cross-price effect of changes in net wages. A difference is that the fiscal externality terms are multiplied by N. This captures the fact that the implied increase in the effective marginal tax rate is larger the greater the household demand for public childcare. 2(a) and 2(b) reflect responses to changes in net prices $\mathcal{K}^{net}$, which increase for households with $y > y_p$ and decreases for those with $y < y_p$. The own-price

effect of these changes, $O_{\mathcal{K}}$, reflected in 2(a), is a decrease in nursery demand for richer households and an increase in nursery demand for the less well-off. The net impact of these two changes on the government budget and therefore on the MECR is ambiguous. Finally, 2(b) captures the cross-price effect of changes in $\mathcal{K}^{net}$ on the labor supply of all households. Again here we have two effects of opposite signs for households above and below the p-th percentile and the net effect is ambiguous. Finally, note that the changes in the net cost that trigger effects described 2(a) and 2(b) are larger the further away the households are from y_p.

Under what conditions is $MECR(\hat{s}_p) > MECR(\widehat{T}_p)$? Since the net wage effects 1(a) and 1(b) are almost the same for both perturbations, the most important question is whether the net childcare price effects 2(a) and 2(b) add up to an increase or a decrease in the $MECR$. Both 2(a) and 2(b) yield an ambiguous contribution to the MECR since the effects on the populations below and above the p-th percentile are of opposite signs.

The cross-price effect of a childcare price change on labor supply, embodied in point 2(b), depends on the complementarity of nursery use and labor supply for different households. Richer households facing more expensive nursery fees will decrease their labor supply while poorer households are likely to increase their labor supply in the face of cheaper nursery use. For given values of the cross-price elasticity, the magnitude of the fiscal externality $|X_{\mathcal{K}}|$ is increasing in income and the labor wedge τ^H. Therefore, unless low-income households have significantly larger elasticities, the negative fiscal effect of households with income above y_p dominates the positive fiscal effect of households with below y_p income, since not only is income larger above the median, but the labor wedges is as well due to increasing marginal tax rates of the baseline tax-transfer schedule.

Similarly, the own-price effect on the use of nursery services, embodied in point 2(a), depends on how sensitive households with income above versus below y_p are to childcare prices and how much childcare subsidies per hour vary with income.

Which of these effects dominates in the real world is a non-trivial question since the labor and childcare wedges and the various elasticities all vary across the income distribution. There is no direct empirical evidence for the quantification of the terms in Propositions 1 and 2. We therefore develop and estimate a dynamic structural model based on dynamic decision making observed in panel data. We will relate our quantitative findings to Propositions 1 and 2 in a transparent manner. Before we turn to our structural quantitative analysis, we now discuss how the comparison of $MECR$s is further affected by dynamic effects such as future maternal earnings and child development considerations.

2.5 Dynamic effects

2.5.1 Parental decisions

So far, we considered a static version of the model. The logic simply extends to a dynamic setting. If there were more than one period, the perturbations $\widehat{T}_p$ and $\hat{s}_p$ may affect future earnings or childcare demand of parents even if these changes are only implemented in one period. An important channel is the dynamics of mothers' wages: less labor supply today results in lower wages and lower tax payments in the future. Incorporating this into the analysis is straightforward. As opposed to considering a static budget neutrality, one can consider a net-present value dynamic budget neutrality. These changes in mothers' future earnings then affect the MECR simply through their impact on $\theta^b(\theta^a)$ and $\sigma^b(\sigma^a)$. The fiscal externalities will get enlarged by a dynamic component. This will be included in our structural model.

2.5.2 Child development

We now show how we integrate the impact of nursery attendance on children's future earnings into the MECR analysis. For this purpose, denote the net-present value of lifetime earnings of a child from household i as y_i^{NPV}. Further, denote the change in this net-present value of lifetime earnings due to a marginal increase in N_i as

$$dy_i^{NPV} = \frac{\partial y_i^{NPV}}{\partial N_i}.$$

Generally, dy_i^{NPV} can be < 0 or > 0. As we discuss below in our quantitative section in detail, the empirical evidence points $dy_i^{NPV} > 0$ for low-income households and $dy_i^{NPV} \leq 0$ for children from high-income households.

Fiscal externalities. Changes in children's earnings caused by the perturbations under study imply fiscal externalities. Formally, the fiscal externality through cross-price effects of wages on childcare demand (14) becomes:

$$X_w^i = -\frac{\tau_i^N + \tau_i^{npv} dy_i^{NPV}}{1 - \tau_i^H} \cdot N_i \cdot \varepsilon^i_{N,w_f^{net}} \tag{14}$$

and the fiscal externality through an own-price effect of the net childcare cost on childcare demand becomes:

$$O_{\mathcal{K}}^i = -\frac{1}{\mathcal{K}} \cdot \frac{\tau_i^N + \tau_i^{npv} dy_i^{NPV}}{1 - \tau_i^N} \cdot N_i \cdot |\varepsilon^i_{N,\mathcal{K}^{net}}| \tag{15}$$

where τ_i^{npv} is the effective maginal tax rate on lifetime incomes of children of household i and therefore $\tau_i^{npv} dy_i^{NPV}$ captures the fiscal effect of a marginal change in N_i through its

impact on the child's earnings.

These changes in X_w^i and $O_{\mathcal{K}}^i$ affect $\theta^b(\theta^a)$ and $\sigma^b(\sigma^a)$ which in turn affects the MECRs as defined in (8) and (9) For example, if low-income children benefit from attending public childcare in terms of their lifetime earnings, the MECR of childcare subsidies will be lower, *ceteris paribus*, since the childcare fee change increases their attendance at a nursery. On the contrary, the MECR of the tax-transfer change *ceteris paribus* will be increased because this perturbation lowers their nursery attendance.

Compensating variations of children. Children themselves are also directly affected by the implied changes in their earnings. We assume that the compensating variations of the children induced by these earnings changes is given by:

$$CV_{child,i} = (1 - \tau_i^{npv}) dy_i^{NPV}.$$

Next, we define $MECR^c$, which accounts for children's earnings being endogenous to today's nursery attendance.

$$MECR^c(\widehat{T}_p, \rho^c) = 1 - \frac{CV^b_{parent}\left(\widehat{T}_p\right) + \rho^c \; CV^b_{child}\left(\widehat{T}_p\right)}{\left| \; CV^a_{parent}\left(\widehat{T}_p\right) + \rho^c \; CV^a_{child}\left(\widehat{T}_p\right) \; \right|} \tag{16}$$

$$MECR^c(\hat{s}_p, \rho^c) = 1 - \frac{CV^b_{parent}\left(\hat{s}_p\right) + \rho^c \; CV^b_{child}\left(\hat{s}_p\right)}{\left| \; CV^a_{parent}\left(\hat{s}_p\right) + \rho^c \; CV^a_{child}\left(\hat{s}_p\right) \; \right|} \tag{17}$$

where ρ^c denotes the Pareto weight the social planner puts on children relative to the generation of their parents. The CV of the parents are as in Definition 1, the subscript "parent" is only added for clarity. Note that $\rho^c < 1$ would not capture standard discounting of future payoffs. The latter is already implicit in the definition of $CV_{child,i}$. Instead $\rho^c < 1$ would capture that the social planner assigns less weight on future generations. The MVPF literature typically sets $\rho^c = 1$ (Hendren and Sprung-Keyser 2020), implying equal weights for all generations. This approach aligns with the intergenerational equity perspective of Stern (2008), which advocates for giving future generations the same moral consideration as the present. While Stern's perspective aggregates utilities and not compensating variations, which leads to differences when marginal utilities of consumption differ across generations, $\rho^c = 1$ remains a pragmatic benchmark that reflects the spirit of Stern's equity framework.

For the case where parents are altruistic, our approach effectively accounts for both independent moral significance and intergenerational altruism (as in Farhi and Werning (2010), see Eden (2024)). As noted by Eden (2024), this "double counting" of children's welfare is negligible for social rankings when extending the analysis to include multiple generations (past, present, and future). Thus, normative implications remain robust even when this dual consideration is present.

In our quantitative analysis below, we consider alternative values of ρ^c, with $\rho^c = 1$ as

our preferred benchmark.[6]

3 Dynamic structural model

As shown in Section 2, the magnitude of the MECR of both policy instruments depends on the distribution of households' elasticities of labor supply and public childcare demand with respect to w^{net} and $\mathcal{K}^{net}$. Our structural model adapts the model in Turon (2019) and aims to provide a rich picture of households faced with labor supply and childcare choices within a dynamic framework. It distinguishes itself from the model in Turon (2019) with the addition of a third use of time for the mother besides labor supply and childcare, namely leisure, and a third mode of childcare besides public childcare and maternal care, namely informal childcare provided by relatives or friends. Both the preference for leisure and the availability of informal childcare are allowed to vary across households.

Our environment is composed of households with two adults with up to three children. Households' decision making is unitary and forward looking. The unit time period is 3 years. Marriages are formed at the age of 20 and are stable. Both spouses are of the same age, retire at 65, and have a remaining lifespan of 15 years after retirement. Fertility follows an exogenous stochastic process, which captures the substantial empirical heterogeneity in family composition and in the age of parents at first birth.

Households with young children make two decisions each period: how to provide care for their children and how much maternal labor to supply. Regarding childcare, they decide between the mother caring for the children at home, which we call 'domestic childcare', and externally provided childcare. The latter can either be informal childcare by, e.g., grandparents, or the use of public childcare services, which we call 'nursery'.[7] labor supply choices are discrete: The female spouse can work full-time, part-time, or choose not to participate, while the male spouse is assumed to always work full-time.[8] An important dynamic component of our framework comes from the positive impact of current working hours on the expected growth rate of future wages.

A distinct feature of our model is the large amount of heterogeneity. Households differ in education, which is an important component in the stochastic wage and fertility processes, and in the number and ages of the children they currently have. Besides education, female wages, male wages and children demographics, households are heterogeneous in three

[6]In addition to the result by Eden (2024) that altruism is irrelevant for social discounting, it may be the case that some parents are not altruistic. And even if they are, one could question the application of the envelope theorem for their public childcare decision. Empirical research finds that parents – in particular low-income parents – underestimate the gains from early childhood education (Boneva and Rauh 2018, Cunha, Elo, and Culhane 2022). Correcting such wrong beliefs would require to adjust even the compensating variation of the parents themselves to account for the positive benefits.

[7]The introduction of informal childcare is motivated by the fact that we observe some mothers who work more hours than they buy public childcare for, see Figure B.1.

[8]Close to 90% of fathers of children below 9 work full-time in our sample. Therefore we rule out that fathers provide domestic childcare during working hours.

further (unobserved) dimensions: their preference for domestic childcare, their taste for the female spouse's leisure, and their access to free informal childcare. As we argue below, accounting for this unobserved heterogeneity is key to capturing the large heterogeneity in childcare and labor supply choices of households.

3.1 Children

Children are born, one by one, to parents between the ages of 20 and 40. Subsequent siblings can only be born one or two 3-year interval(s) later, i.e., all age gaps between children of a family can only be 3 or 6 years. The fertility process is stochastic and is determined by the education and the age of the mother and the presence of older siblings.

For our model purposes, the child age ranges that are relevant are (0 – 2), (3 – 5), (6 – 8), and (9+). We denote χ a 4-element vector indicating the presence of a child in each of these age brackets. For example, a family whose composition is represented by the vector $\chi = (0, 1, 1, 0)$ has two children, the youngest aged between 3 and 5 and the eldest aged between 6 and 8. By assumption, each of the first three elements of χ can only be 0 or 1 since only one child can be born in each period. Transitions between different values of χ are governed by (stochastic) fertility events and the (deterministic) ageing of the children. Finally, we assume that households cannot have more than three children.[9]

3.2 Preferences

As in our simple model presented in Section 2, households value female leisure time L, household consumption C, and domestic childcare D. Household consumption is made comparable across different household sizes k by applying a square root equivalence scale. Preferences are reflected in the following instantaneous utility function:

$$u(C, L, D) = \Big(1 - \mathcal{G}(g, \chi)\Big) \left((1-\alpha) \, \frac{\left(\frac{C}{\sqrt{k}}\right)^{1-\gamma_c} - 1}{1-\gamma_c} + \alpha \, \frac{L^{1-\gamma_L} - 1}{1-\gamma_L} \right) + \mathcal{G}(g, \chi) \, \frac{D^{1-\gamma_D} - 1}{1-\gamma_D},$$

where the three CRRA coefficients γ_c, γ_L and γ_D are homogeneous across all households.

Preference heterogeneity. Households' preferences are heterogeneous in two dimensions: α represents the relative taste for female leisure over consumption, and g is the relative preference for domestic childcare. We allow the taste for domestic childcare to vary with the age of the child by introducing $\mathcal{G}$ as follows:

$$\mathcal{G}(g, \chi) = \begin{cases} g & \text{if youngest child's age} \in [0, 3), \\ g \cdot \kappa & \text{if youngest child's age} \in [3, 9). \end{cases}$$

[9] Only 4.99% of households have more than three children in our data.

This allows us to capture the sharp difference in public childcare enrolment between 0 – 2 and 3 – 5 year old children, see Section 5.2. While we are introducing g as a preference parameter, we think of g in a more general sense as a reduced form which may capture i) the true preference heterogeneity for spending time with the child; ii) heterogeneity in how much parents (dis)like their child being in nursery or informal childcare, e.g., due to social norms or trust in the quality of the childcare institutions; iii) the fixed utility cost of bringing children to nursery.[10] It does not necessarily internalise the long-term implications of the mode of childcare on long-term children's outcomes since parents, as indeed the academic literature, only have an imperfect estimate of those returns.

3.3 Constraints

Childcare hours constraints and childcare expenditures. We now describe the time constraint for childcare provision. For each $j = 1, 2, 3$ relating to the age ranges (0 – 2), (3 – 5), and (6 – 8), a child needs an age-specific number of hours of childcare, $\bar{t}_j$, within normal working hours (40 hours per week). In the first and second child age categories, the child needs care all of the time, whereas in the third category, the child needs care in the non-school hours only since she is enrolled in compulsory primary school already. Apart from domestic childcare, households may fulfil the childcare need by calling on informal childcare providers, e.g., grandparents, denoted I, or public childcare services, i.e. a nursery, denoted N. Informal childcare is free and only available to some households. The variable I ranges between 0 and 40 hours a week. If available, households always prefer to use I hours of costless informal childcare over N hours of costly public childcare. In that sense, I captures both whether informal childcare is available and if it is considered equally good as public childcare. Public childcare is always available at an hourly fee of $p\,(j, \chi, y)$, which depends on the age j of the child, the family structure χ, and the household gross income y. In terms of our simple model in Section 2, p corresponds to $\mathcal{K} - s_i(y)$, i.e. the subsidized price of childcare. For a given amount of domestic childcare and informal childcare use, the resulting amount of public childcare necessary for a child of age j is thus given by:

$$N(j) = \max\left\{0, \bar{t}_j - D - I\right\} \tag{18}$$

We also define the share of childcare needs that a household covers with public childcare:

$$m(\chi) = \frac{\sum_{j=1}^{3}\ \chi(j) \cdot N(j)}{\sum_{j=1}^{3}\ \chi(j) \cdot \bar{t}_j} \tag{19}$$

where, as defined above, $\chi(j)$ is the j-th element of the vector χ indicating if a child of age j currently lives in the household. The household expenditure on public childcare of

[10] Distance to the childcare facility could be one potential reason why parents do not send their children to public childcare.

all its children is thus given by S:

$$S(\chi, y) = \sum_{j=1}^{3} \chi(j) \cdot N(j) \cdot p(j, \chi, y).$$

This equation clearly shows that public childcare implies higher expenditure the more children a household has because childcare fees have to be paid for all children. This contrasts with domestic or informal childcare, where the same unit of time can be used to look after one to three children.

Parental time constraint. At each age t, the household has to choose between female labor supply (H_t), female leisure (L_t), and the provision of domestic childcare (D_t). Hence the time constraint is written as:

$$H_t + L_t + D_t = 40. \tag{20}$$

where 40 captures the usual weekly full-time hours.

Budget constraint. We abstract from borrowing and saving to keep the state space tractable despite the large amount of heterogeneity. In that sense, the budget constraints are static and given by:

$$C_t + S(\chi_t, y_t) = y_t - T(y_t, \chi_t), \tag{21}$$

where

$$y_t = 40 \cdot \ w_{m,t}\left(w_{m,t-1}\right) + H_t \cdot w_{f,t}\left(w_{f,t-1}, H_{t-1}\right).$$

$T(y_t, \chi_t)$ captures the tax-transfer system depending on household income as well as the number and age-composition of the children. $H_t \in \{0, 20, 40\}$ represents non-participation, part-time, and full-time work respectively. Male wages $w_{m,t}$ and female wages $w_{m,t}$ are assumed to follow first-order Markov processes. For women, transition rates between wage grid points depend on current labor supply H_t. Finally, once the spouses retire, they get a fraction $\mathcal{B}$ of their last period's full-time earnings potential as retirement benefits.

3.4 Dynamic decision problem

We summarise all heterogeneity in the following vector:

$$\Omega_t = (s_t, h) \quad \text{with} \quad s_t = (t, w_{m,t}, w_{f,t}, \chi_t, educ) \quad \text{and} \quad h = (g, I, \alpha)$$

At each age t, the household chooses female labor supply (H_t) and the amount of domestic childcare (D_t). These imply the values of consumption (C_t), female leisure (L_t), and the use of public childcare (N_t). The three constraints that the household faces are the need for childcare (18), the time constraint for the female spouse (20), and the budget constraint

(21). The dynamic household problem is defined for a given state space vector Ω_t as:

$$V(\Omega_t) = \max_{H_t, D_t} u(C_t, L_t, D_t|\Omega_t) + \beta \mathbb{E}[V(\Omega_{t+1}|\Omega_t, H_t)], \text{ s.t. (18), (20) and (21).} \quad (22)$$

The model is solved by backward induction from retirement. We assume that during retirement all income is used as consumption and the entire time endowed per period is used as leisure since, by construction, there are no children to be taken care of anymore.

The distribution of (g, I, α) conditional on observables is key to capturing the observed behavior of households. First, we allow for heterogeneity in leisure preferences α to account for the sizeable variation in labor supply conditional on wages. Such heterogeneity in leisure preferences (or equivalently, disutility of work) is a common component in structural models to match hours worked (such as in, e.g., Blundell et al. 2016). A more distinctive feature of our model is the heterogeneity in g and I. It is necessary to account for the heterogeneity in childcare decisions conditional on observables that we observe in the data.

4 Estimation methodology

Our estimation procedure can be decomposed into two parts. First, in Section 4.1, we estimate and calibrate various parameters without using the explicit structure of the model. In a second step, in Section 4.2, we quantify the remaining parameters by using the model structure.

4.1 Auxiliary regressions

4.1.1 Policies and childcare need

In this section, we first calibrate the childcare need of the different age groups. We also calibrate the costs for the government of providing a full-day public childcare slot and estimate the childcare fee schedule as a function of parental income. Finally, we calibrate the government policies required as exogenous inputs for our model.

Childcare need. The age-specific weekly hours of childcare needed, $\bar{t}_j$, are calibrated as follows, for each child age j. If a child is younger than 6, the childcare need is set to 40 hours per week, i.e., 100% of the usual working week. To account for the fact that nearly all 3 – 5-year olds attend kindergarten at least half-days , we impose that 20 of the possible 40 hours for this age group have to be covered by public childcare. For children aged 6 – 8, the need reduces to 15 hours per week because these children attend compulsory schooling for 25 hours per week.

Public childcare cost structures. We approximate the cost structure of public childcare institutions by assuming the costs to the government to be linear in the number

of children. We use the values in Table 1, provided by the German Statistical Office.

Table 1: Average annual cost per child for 40h/week of public childcare

Children's age interval	0 – 2	3 – 5	6 – 8
Annual cost	€11,837	€7,927	€6,733

Notes: See Statistisches Bundesamt (2012), converted to 2017 prices.

Childcare fees. The 2013, 2015, and 2017 waves of our GSOEP data contain information on public childcare hours per day and monthly fees paid. We use the same sample restrictions as in our structural sample that we describe below in Section 4.2.1. In Appendix C.1.2 we estimate childcare fees as a function of gross household income, which we interact with the number of siblings. Hourly fees are found to increase with income and decrease with the number of siblings (per child). E.g., for a full-time slot for a 0-2 year-old child, the marginal price is 3%: for a 100 Euro increase in earnings, the monthly full-time childcare fee increases by 3 Euro.[11]

Taxes. We use the Matlab implementation of the German tax and transfer code provided by Bick et al. (2019) to map gross to net income and calculate tax revenues. The implementation is based on the annual OECD "Taxing Wages" reports and takes into account federal income taxes as well as social security contributions, cash benefits, and standard deductions.[12]

Pensions. We approximate the German pension system by assuming that households receive 40% of both partners' last period's potential gross full-time earnings (OECD 2017).

Interest rate. We set the real interest rate of the government to 6% per 3-year model period, which corresponds approximately to 2% per annum.

4.1.2 Estimation of the fertility process

We estimate the fertility process in Appendix C.2 as transition probabilities consistent with our model assumptions set out in Section 3.1. We use the 2014 and 2018 Microcensus

[11]Since subsidies also vary between regions, we ran these regressions with state dummies and dummies for living in an urban region as a robustness check. The slope coefficients on income were very similar.

[12]When calculating the fiscal effects of changes in labor supply, we account for the sum of income tax payments, social security contributions for public sickness and care insurance, and solidarity surcharge payments. We disregard social-security spending because the German Bismarckian pension system implies pension benefits that are proportional to social security contributions paid; there is no concavity in the benefit formula as, for example, in the U.S. Aside from a precise implementation of the non-linearities of the tax code, it includes joint taxation of couples as well as child benefits for each child in the household. Marginal tax rates faced by women vary with their spouses' income and child allowances reduce the taxable income of the household.

waves and focus on births taking place from 2012 to 2017 (see Appendix C.2 for more details on the sample). Figure 2 illustrates our estimates in terms of the evolution of shares of families with zero to three children over the age of the mother and by education level, referring to having obtained an A-level or not. In terms of completed fertility, the figures are similar in both education groups: about 45% of households have two children, about 30% (respectively 10%) have one (respectively three) child(ren) and about 15% of households remain childless. The timing of births, however, differs markedly between education levels, with low-educated women having children earlier. By age 34 (respectively 37) for the low (respectively high) education group, the majority of households have completed their fertility.

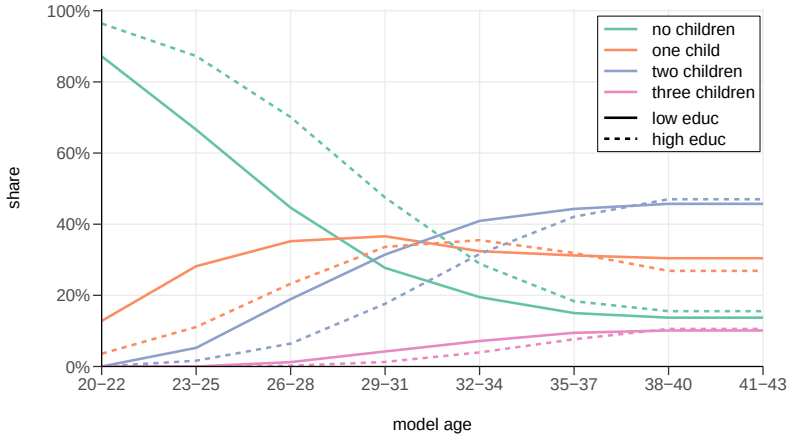


Figure 2: Family composition as implied by the fertility process

Notes: 'low educ' corresponds to no A-level, 'high educ' to having obtained an A-level. Sources: FDZ-StABL (2020a), FDZ-StABL (2020b).

4.1.3 Estimation of the wage process

We estimate the following equation for the wage process of women:

$$log(w_{f,it}) = \alpha + \beta_1 log(w_{f,it-1}) + \beta_2 \mathbb{1}\{lm_{it-1} = NP\} + \\ \beta_3 \mathbb{1}\{lm_{it-1} = PT\} + \beta_4 educ_i + \mathcal{A}(t) + \varepsilon_{it}^{w_f},$$

where $\mathbb{1}\{lm_{it-1} = NP\}$ and $\mathbb{1}\{lm_{it-1} = PT\}$ are dummy variables that indicate whether a woman i was either not working or working part-time in period $t-1$ and $\mathcal{A}(t)$ is a third-order polynomial in age.[13] The GSOEP sample is an extended version of our structural sample that we describe in detail in Section 4.2.1. To increase the power of this

[13]We have omitted here the selection term, but we describe in Appendix C.3.1 the detail of our joint estimation of wages and participation into work *à la* Heckman.

regression, we consider a larger time span, namely 2000-2017, see Appendix C.3.2.

The estimated wage penalties for working part-time or not working instead of working full-time are substantial and amount to 5.5% and 16.5% per 3-year model period.[14] In

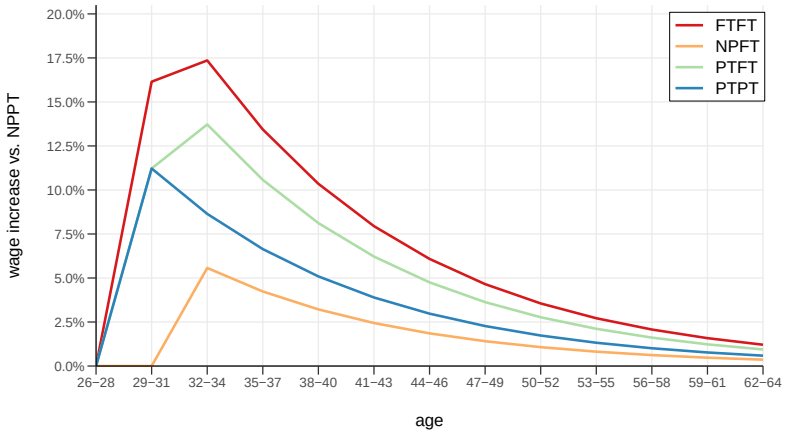


Figure 3: Illustration of the female wage process

Notes: Relative increase in wages of different labor supply patterns, always compared to not working at age 26 – 28 and working part-time at age 29 – 31. NP, PT, and FT denote not working, part-time work, and full-time work, respectively. PTFT denotes part-time work at age 26 – 28 and full-time work at age 29 – 31. All other patterns are defined analogously. Simulations based on female wage process estimates from Appendix-Table C.2.

Figure 3, we illustrate our wage process estimates by looking at the benefits of increasing labor supply relative to a typical labor supply pattern of mothers. Specifically, we consider a mother who has her first child at 26. The benchmark is that she does not work while the child is 0 – 2, works part-time when the child is 3 – 5, and works full-time afterwards. The graph illustrates the dynamic wage gains that the mother would obtain if she increased her labor supply relative to the benchmark. The blue line shows the case when the mother already starts working part-time when the child is 0 – 2. The red line shows the case where the mother switches to full-time work both when the child is 0 – 2 and when the child is 3 – 5. Aside from the substantial wage gains from increasing labor supply, the graph clearly illustrates that the potential wage gains are quite persistent. Finally, we also estimate the male wage process in a similar fashion (see Appendix-Table C.2), but without part-time or non-employment penalties since we focus on full-time working males.

[14] See Appendix-Table C.2. The numbers are consistent with evidence by Ejrnæs and Kunze (2013).

4.2 Structural estimation

4.2.1 Estimation sample

We focus on German mothers aged between 20 and 65 who are currently not in education and share a household with a full-time working partner. We track this group over the time span 2012 to 2017 in a representative longitudinal survey data set, the German Socio-Economic Panel (GSOEP). We allocate all children into the corresponding model age brackets and the household into the corresponding child-age structure χ. We only keep households with complete information for two model periods and average all household variables of interest within each assigned period. This leaves us with an estimation sample of 2,182 households. 1,076 of these households face some childcare needs in at least one of the two periods. The other half of the sample does not face childcare needs in either period as their children are aged 9 or older in both periods. Nevertheless, we keep these in the estimation sample as they help to identify heterogeneity in leisure preferences α. Further details on the data and the assignment procedure can be found in Appendix D.1 and summary statistics on the sample are presented in Table 2. To understand the role of last four rows, we refer the reader to the paragraph 'Constant characteristics x' in Section 4.2.3.

Table 2: Summary statistics for the MLE sample

	mothers of 0 – 8 year olds	mothers of 9+ year olds
age	36.10	50.54
w^{male}	23.56	23.94
w^{female}	17.26	16.16
share high education	51%	32%
number of 0-9 children	1.28	
age of youngest child	3.41	
share living in former East	20%	27%
share demanding occupation	43%	36%
share catholic	29%	32%
share urban	64%	60%
N	1,076	1,106

Notes: *east*: lived in former East Germany at some point; *demanding occupation*: primarily worked in an occupation where at least one-third of the tasks can be classified as analytic non-routine; *catholic*: being Catholic at age 20; *high education*: at least an A-level; *urban*: living predominantly in an urban area. Sample: mothers aged 20 to 65 who are not in education and live with a full-time working partner. Source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).

4.2.2 Parameters set externally

In line with Blundell et al. (2016), we set the discount factor β to 0.94 per 3-year model period. For γ_c we follow common practice in macroeconomics and assume that consumption enters the utility function in a logarithmic manner (Bick and Fuchs-Schündeln 2018, Guner, Kaygusuz, and Ventura 2020). We also use a logarithmic functional for domestic childcare, i.e. set $\gamma_D = 1$. To avoid Inada conditions we set a floor value of 1 weekly hour for both L and D. This leaves us with two further parameters to set: γ_L and κ. We calibrate these parameters in the following fashion: we chose the value of γ_L to obtain a Hicksian intensive margin elasticity that comes close to the value of 0.33 that is widely used in the public finance literature and goes back to Chetty (2012). Furthermore, the shifter for the preference for domestic childcare κ is set such that the model predicts well how childcare demand differs by the age of the youngest child in the family.[15] The value of 0.075 implies that households have a much stronger preference to spend time with children below age 3 compared to 3 to 8 year old children. Besides having a stronger preference to be with the child while it is young, this could also be considered as a reduced form for social norms.[16]

Table 3: Parameters set externally

parameter	β	γ_c	γ_D	γ_L	$\overline{L}$	$\overline{D}$	κ
value	0.94	1	1	1.75	1	1	0.075

4.2.3 Maximum likelihood estimation of heterogeneous preferences

Our data comprises observations of female labor supply H_p and total public childcare take-up of the household $m_p^h(\chi_p)$ for two model periods $p = 1, 2$. We estimate the distributions of unobserved heterogeneities $h = (g, I, \alpha)$ to maximize the likelihood of these dynamic choices. In practice, we estimate distributions of (g, I, α) conditional on observables x, which we introduce below. As a first step, we now build up the likelihood function step by step and start with measurement-error components.

Measurement Error. We allow for some measurement error in the wages of both spouses and in the amount of public childcare consumed. We denote observed wages and total public childcare as $(\widetilde{w}, \widetilde{N^h})$, in contrast to the 'true' quantities (w, N^h).[17] Since it is measured in a discrete manner, the labor supply of mothers is assumed to be error-free.

[15] For computational reasons we did not include the estimation of κ into our maximum likelihood estimation that we describe below. Instead we conducted the MLE conditional on different values of κ. We iterated over several values and picked the one which implied the best model fit in terms of childcare.

[16] In the 2016 wave of the German General Social Survey around 40% of respondents agree with the statement "A small child is bound to suffer if his or her mother goes out to work." Source: GESIS (2017).

[17] For our model, we discretise weekly hours of public childcare in 2.5 hours steps, i.e. $N \in \{0, 2.5, 5, ..., 40\}$. We assign the values observed in the data to these discretised values.

We denote the errors as follows:

$$\begin{aligned}\log(\widetilde{w_{p,q}}) &= \log(w_{p,q}) + \epsilon_{p,q} \text{ with } |\epsilon_{p,q}| \sim EV_{\text{II}} \\ \widetilde{N_p^h} &= N_p^h + u_p \text{ with } |u_p| \sim EV_{\text{II}}\end{aligned}$$

for $q = f, m$ and $p = 1, 2$. All measurement errors are assumed to be distributed as type II extreme value distributions.[18] We further assume that they are independent so that the joint distribution of errors is:

$$f(\epsilon_{p,f}, \epsilon_{p,m}, u_p) = ev_w(\epsilon_{p,f}) \cdot ev_w(\epsilon_{p,m}) \cdot ev_T(u_p)$$

where $ev_d(.)$, for $d = w, T$ denotes the density of the type II extreme value distribution for wages and childcare hours respectively. The likelihood of observing the choices of a household in period p can then be written as:

$$\ell(H_p, \widetilde{N_p^h} | \widetilde{s_p}, h) = \int\int\int \ell(H_p, N_p^h | s_p, h) \cdot f(\epsilon_{p,f}, \epsilon_{p,m}, u_p) \cdot d\epsilon_{p,m} d\epsilon_{p,f} du_p.$$

where $\widetilde{s_p}$ denotes the time varying state space including the observed wages $\widetilde{w_{p,f}}$ and $\widetilde{w_{p,m}}$. Note that for this intermediary step, we condition on unobservables h.[19]

The likelihood of the 'true' choices (H_p, N_p^h) matching the model predictions $\widehat{H_p}$ and $\widehat{N_p^h}$ for a household with characteristics (s_p, h) is:

$$\ell(H_p, N_p^h | s_p, h) = \begin{cases} 1 & \textit{iff} \quad \widehat{H_p}(s_p, h) = H_p \text{ and } \widehat{N_p^h}(s_p, h) = N_p^h, \\ 0 & \text{otherwise.} \end{cases}$$

The likelihood of observing a household's sequence of choices $(H, \widetilde{N^h})$ given the full set of time-varying characteristics $\widetilde{s} = (\widetilde{s_1}, \widetilde{s_2})$ and unobserved heterogeneity h is thus:

$$\ell(H, \widetilde{N^h} | \widetilde{s}, h) = \prod_{p=1}^{2} \ell(H_p, \widetilde{N_p^h} | \widetilde{s_p}, h).$$

Our object of interest is the joint distribution of unobserved heterogeneity $\ell(h|x)$ conditional on a set of constant household characteristics, denoted x. The likelihood of observing a household's sequence of choices $(H, \widetilde{N^h})$ conditional on observed characteristics

[18] For the measurement error in wages, we set the scale and shape parameters to $\sigma = 0.026$ and $\xi = 0.5$ to ensure that 90% (respectively 95%) of errors are no more than 20% (respectively 40%). This is in line with the literature (e.g., Blundell et al. (2016)). The calibration of the measurement error in nursery hours is such that 90% of the errors are no more than 5.6 hours ($\sigma = 1.05$ and $\xi = 0.5$).

[19] This specification is assuming that wages are only correlated with h via the impact of h on labor market choices. In other words, conditional on labor market choices, wages are independent of h.

is given by the following expression:

$$\ell(H, \widetilde{N^h}|\widetilde{s}, x) = \int_{\boldsymbol{h}} \ell(H, \widetilde{N^h}|\widetilde{s}, h) \cdot \ell(h|x)\ dh.$$

Finally, our sample likelihood is the product of all individual likelihood contributions of the N households in our data:

$$\mathcal{L} = \prod_{n=1}^{N} \ell(H^n, \widetilde{N}^{h,n}|\widetilde{s^n}, x^n). \tag{23}$$

Joint distribution of unobserved heterogeneity. We now zoom into the joint distribution of unobserved heterogeneity $\ell(h|x)$. We assume the marginal distributions of g, I, and α, to be independent conditional on constant characteristics x:

$$\ell(\underbrace{g, I, \alpha}_{=h}|x) = \ell^g(g|\boldsymbol{x}^g) \cdot \ell^I(I|\boldsymbol{x}^I) \cdot \ell^\alpha(\alpha|\boldsymbol{x}^\alpha),$$

where $\boldsymbol{x}^g, \boldsymbol{x}^I$, and $\boldsymbol{x}^\alpha$ are subsets of x that are allowed to overlap. Such overlap creates correlations between the marginal distributions ℓ^g, ℓ^I, and ℓ^α without assuming an explicit correlational structure. We assume each type of heterogeneity, $het \in \{g, I, \alpha\}$, to be normally distributed conditional on $\boldsymbol{x}^{het}$:[20]

$$het^\mu|\boldsymbol{x}^{het} \sim \mathcal{N}\left(\gamma^{het} + \boldsymbol{x}^{het}\boldsymbol{\beta}^{het}, 1\right).$$

Our maximum likelihood procedure will estimate the parameters (γ^g, $\boldsymbol{\beta}^g$, , γ^I, $\boldsymbol{\beta}^I$, γ^α, $\boldsymbol{\beta}^\alpha$) which maximize the sample likelihood function given in equation (23).

Constant characteristics x. The fact that we allow the joint distribution of unobserved heterogeneity h to vary with characteristics x is important in multiple ways: The characteristics in x allow us to capture that subgroups in our data may have very different preferences, thereby improving the capability of our model to predict the behavioral patterns in the data. These constant characteristics also help us address the initial conditions problem, i.e., that the time-invariant joint distribution of unobserved heterogeneity might have affected the initial values of our time-varying state variables. Summary statistics of the variables in x can be found in Table 2.

The variables selected in $\boldsymbol{x}^g$ are indicators variables for living in East Germany, for being Catholic at age 20, and for having primarily worked in a 'demanding occupation'.[21]

[20]The heterogeneity terms g and α are defined on the closed interval $[0, 1]$, while I is measured in hours and lies on $[0, 40]$. We truncate the normal distribution of $het^\mu|\boldsymbol{x}^{het}$ to the corresponding support.

[21]This is motivated by Adda, Dustmann, and Stevens (2017), who have shown that women select themselves out of analytical jobs if they prefer to spend time with their children. Specifically, we code an occupation as a 'demanding occupation' if the share of interactive non-routine tasks is greater than one-third. We use the task classification of 3-digit occupations by Dengler, Matthes, and Paulus (2014).

The variables selected in $\boldsymbol{x}^I$ are indicators of East/West Germany, maternal education and whether the household lives in an urban area.[22] Finally, the variables included in $\boldsymbol{x}^\alpha$ are maternal education and 'demanding occupation'.

Identification. In the absence of a formal proof, we provide an intuition for the identification of the time-invariant parameters that govern the joint distribution of unobserved heterogeneity. There are three interacting ingredients that identify the time-invariant unobserved heterogeneity: i) cross-sectional variation in choices conditional on observed states, ii) the longitudinal dimension of our panel data, iii) using data for both households with small children and those with older children. Appendix D.2 describes these three ingredients in more detail and argues that our model is credibly identified.

Bootstrap inference To quantify sampling uncertainty, we conduct a household-level bootstrap. In each of $B = 50$ replications, we resample complete two-period household histories with replacement, re-estimate the maximum-likelihood parameters governing the distribution of unobserved heterogeneity $h = (g, I, \alpha)$, and recompute the two dynamically budget-neutral perturbations at the median and their associated MECRs using the re-sampled households. The bootstrap therefore captures both uncertainty in the estimated preference distribution and sampling variation in the empirical distribution of household characteristics and incomes over which the MECRs are evaluated. All inputs calibrated or estimated outside the likelihood are held fixed.[23] See Appendix D.6 for more details.

5 Estimation results

5.1 Results

Since our estimated coefficients do not carry much intuitive meaning, we show them in Appendix D.4, where details on the optimisation routine and on the sensitivity of the estimates are also discussed. We comment here on the main features of our estimates.

First, considering the preference for domestic childcare (g), women who live in former East Germany have lower preferences for domestic childcare than those in former West Germany. Additionally, for Catholic mothers, we observe a higher preference for domestic childcare. Second, we find large differences in the distribution of the availability of informal childcare (I) between East and West Germany. Only very few households in East Germany rely on informal childcare, see also Footnote 22. Focusing on West Germany, the

[22] More specifically, we estimate the distribution of the availability of informal childcare $\boldsymbol{x}^I$ separately for East and West Germany. Social norms about childcare differ significantly between both regions, which likely affects the availability of grandparental childcare. See Hank, Tillmann, and Wagner (2001) for a discussion. For households living in East Germany, we only estimate the intercept.

[23] Ideally we would also estimate the uncertainty around our results coming from the uncertainty regarding the values of the calibrated parameters, but this would be very time consuming to run.

coefficients show that lower educated mothers as well as those not living in urban areas have a higher availability of informal childcare. Third, highly educated women and those having primarily worked in a demanding occupation tend to have a lower preference for leisure (α).

5.2 Model fit

We now turn to the model fit. Specifically, we evaluate the ability of the estimated model to match data moments for the two choices, female labor supply H and household use of public childcare N^h, as a function of the age of the youngest child. As shown in Table 4, overall, we are able to achieve a good fit of the labor supply patterns of mothers conditional on children's ages. In particular, the model matches the observed increase in labor supply once the youngest child turns 3. Part of the increase in labor supply in the data originates from the fact that most mothers take parental leave for one year after child birth. We thus observe in the data a much larger share of public childcare take-up if children are between 12–36 months old which implies that the childcare subsidies are mainly affecting those children. We nonetheless capture both effects by taking three-year averages and hence, combine 0–2 year-old children for tractability of the model.[24] Our model is also able to match the labor supply pattern of mothers with completed fertility, i.e., children aged 9 or older.

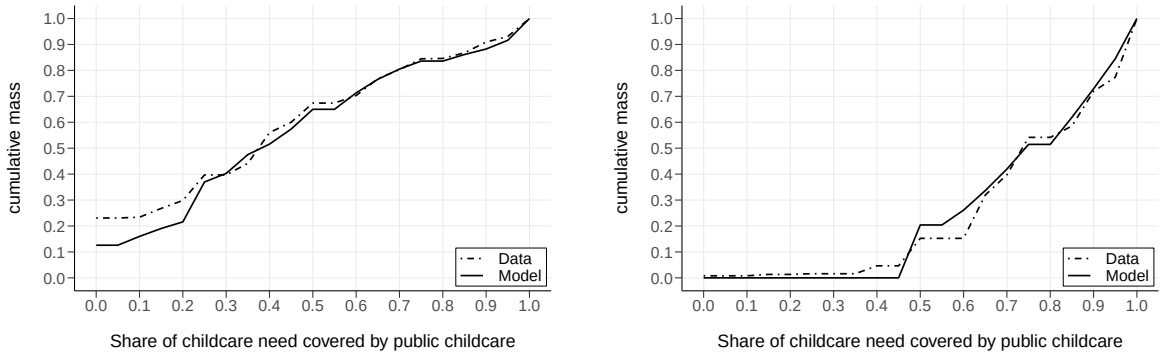
Table 4: Model fit for labor supply by youngest child

	Children 0 – 2			Children 3 – 5		
	NP	PT	FT	NP	PT	FT
Model	0.49	0.44	0.07	0.14	0.64	0.22
Data	0.55	0.40	0.05	0.17	0.65	0.18
	Children 6 – 8			Children 9+		
	NP	PT	FT	NP	PT	FT
Model	0.14	0.64	0.22	0.12	0.64	0.24
Data	0.13	0.68	0.19	0.16	0.54	0.30

Notes: PT and FT denote the female working part-time and full-time, respectively. NP denotes non participation. Sample as defined in Section 4.2.3. Data source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).

Figure 4 illustrates the fit of our model in terms of public childcare demand by the age of the youngest child. Families with the same age of the youngest child might face a different total childcare need as they might or might not have older children. To account

[24] Previous research has shown that maternal time during the first twelve months after child birth is particularly important, see e.g., Carneiro, Løken, and Salvanes (2015). But this particular effect goes beyond the focus of this paper.



(a) Youngest child in the family aged 0 – 2 (b) Youngest child in the family aged 3 – 5

Figure 4: Model fit of families' public childcare demand

Notes: Sample as defined in Section 4.2.3. Data source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).

for this, we plot the share of childcare needs that a household covers with public childcare $m(\chi)$. In general, the model fits the data well. It slightly underpredicts the share of households with zero public childcare if their youngest child is below 3, see Figure 4a. The reason is that our model also allows for quite small amounts of public childcare (starting with 2.5 hours per week) and therefore few households choose a corner solution of 0 hours. Figure E.5 presents the fit for families with the youngest child between 6 and 8.

5.2.1 Model fit by household income

We further illustrate our model fit for households with income levels below and above the median household income calculated from our data sample. Appendix-Table E.6 reports the resulting fit of the labor supply patterns of mothers conditional on the age of their youngest child. The model matches the data very well, capturing the low (high) non-participation and high (low) full-time shares of households with above-median (below-median) income. In addition, Figure E.6 in Appendix E depicts the model fit of public childcare demand for above- and below-median income households. Generally, the model also fits childcare demand well after conditioning on household income and reflects that households with lower income demand less childcare.

5.3 External validity

Our policy experiments in Section 6 below rely on our model predicted responses in terms of female labor supply and public childcare take-up to policy changes. To bring external validity to our policy counterfactuals, we compare the behavioral responses produced by our model with estimates found in the literature.

First, we use our estimates to compute compensated labor supply elasticities at the intensive margin and obtain a value of 0.23. Separating the sample by household income, we find a value of 0.18 for women in below-median income households and 0.30 for above-

median income households. As laid out in Section 4.2.2, we chose the curvature parameter on leisure to obtain values that come as close possible to the literature value of 0.33 (Chetty 2012, Chetty et al. 2011). Given that we constrain the intensive margin response to be only about the part-time vs. full-time margin, it appears reasonable that we stay a little below that number. We also simulate participation elasticities with respect to wage changes and obtain numbers of 0.26 below the median and 0.08 above the median. The average participation elasticity is 0.16. This is in line with the empirical estimates of the participation elasticity found in the quasi-experimental literature and surveyed in Chetty et al. (2011) which are in a range of 0.15 to 0.24 for (non-single) women. Finally, we simulate the propensity to earn out of one unearned Euro. We find that a 1000 Euro increase in unearned income decreases earnings in that period by 213 Euro. This number is higher than the values found in Cesarini et al. (2017) for lottery winners in Sweden and lower than the numbers estimated by Golosov et al. (2021) in the U.S. context.

Second, we consider responses to changes in childcare prices, for which there is much less clear-cut evidence. Gathmann and Sass (2018) and Busse and Gathmann (2020) both provide recent evidence in the German context. Gathmann and Sass (2018) consider the introduction of a so-called homecare subsidy, whereas Busse and Gathmann (2020) use a staggered introduction of free childcare in some German states. We implement such reforms in our model and compare the implied response along the childcare and labor supply margin. As we now argue, our model is broadly in line with their evidence.

We find a decrease in labor force participation of 6.4 percent for the affected population if we introduce the homecare subsidy in our model, which is close to the estimate of 5.1 percent by Gathmann and Sass (2018). Regarding the Busse and Gathmann (2020), they find that, when they control for the number of offered slots, the labor force participation response is 10% but with large standard errors.[25] Our model predicts an increase of 3.4%, which is well within the confidence interval of their estimate.

Turning to responses in childcare demand, we find an 8.3% decrease in childcare use due to the introduction of the home subsidy. This is significantly lower than the 23% reduction in childcare demand from the Gathmann and Sass (2018) estimates. Our model performs much better for the abolition of fees. Busse and Gathmann (2020) find that the introduction of free childcare increases childcare take up by 9.6% for parents with a child between 2 and 3 years. Making childcare free in our model, we find increases of 6.8%. While this figure is smaller than the one found by Busse and Gathmann (2020), we note that our findings below would be accentuated rather then reversed if our prediction was closer to theirs.[26]

[25]We use the estimates from Table 4 of their paper, where they isolate price effects from rationing.

[26]To be precise, we define the public childcare demand extensive margin as the fraction of households increasing their demand from covering less than half of their needs to over half of their needs. The reason is that in our model there are few true zeros because we allow for very small amounts of childcare starting from 2.5 hours per week, see Figure 4a.

6 Quantifying the leaks

We are now in a position to quantify the different components of the MECR for both policy instruments. As discussed in Section 2 these depend on the responses of households across the income distribution to the altered incentives to supply labor and demand childcare services. Our estimates of the structural model presented in Section 5 give us a rich picture of the heterogeneity in these responses, which we now integrate over households below and above the p-th percentile to quantify the leaks of the "redistribution buckets" (8) and (9) and decompose them into the different components described in Propositions 1 and 2.

In practice, we simulate the tax and subsidy perturbations as described in (6) and (7) in our quantitative model. We implement them as temporary changes (i.e., they apply for one three-year model period) and apply the change to the childcare subsidy schedule to childcare for all three child-age groups, $0-2$, $3-5$, and $6-8$. We design them such that they are *dynamically* budget neutral: we account for long-run budgetary effects and take a net-present value perspective on the government budget constraint. For this purpose, we simulate the model until the end of the life cycle for all households in our sample and account for all earnings and childcare demand changes in the future.[27]

In Section 6.1, our benchmark is to conduct these changes around the median household income. In Section 6.2 we take into account child-development effects of public childcare. Finally, in Section 6.3, we show how our model can be used to assess redistributive perturbations around other income percentiles than the median.

6.1 Redistribution from above- to below-median income

Our main finding is as follows: if we abstract from child development, we find $MECR(\widehat{T}_{50}) = 0.28$ and $MECR(\hat{s}_{50}) = 0.42$: for each Euro taken from the top half of the income distribution, 72 Cents reach the bottom of the distribution for the tax-transfer perturbation, and 58 Cents do for the childcare subsidy perturbation. Hence, the marginal efficiency cost of redistribution is 50% larger for the childcare subsidy schedule than for the tax-transfer schedule. The boostrap analysis described in Sectionsec:MLE reveals following bootstrapped standard errors for these results: 0.01 and 0.02 for $MECR(\hat{T}_{50})$ and $MECR(\hat{s}_{50})$, respectively. More importantly, the paired standard error of their difference is 0.02, indicating that the benchmark ranking is not driven by sampling variation in the household sample. More details on this analysis, can be found in Appendix D.6.

To understand the sources of the leakages, we decompose the MECR in Table 5.

The first row captures the leaks which are due to changes in labor supply. For the tax-transfer perturbation, we find that this number is 0.34 – with 0.21 (respectively 0.13)

[27] Note that we are not considering a balanced cohort. In our policy simulation, we include all households who have a 0-8 years-old child between 2012-2017, i.e. those in the left column of Table 2. We then implement our perturbation for the period 2012-2014.

Table 5: Decomposition of MECR

MECR Decomposition	**Tax-transfer system**			**Progressive subsidies**		
	$> y_{50}$	$\leq y_{50}$	$\sum$	$> y_{50}$	$\leq y_{50}$	$\sum$
labor Supply	0.21	0.13	0.34	0.43	0.05	0.47
Childcare Demand	-0.04	-0.03	-0.06	-0.14	0.09	-0.05
Total	0.17	0.10	**0.28**	0.28	0.14	**0.42**

coming from labor supply effects above (respectively below) the median. This quantifies the own-price effect of the change in the net wage O_w, see part 1(a) of Proposition 1. The steeper tax-transfer schedule acts as a disincentive to supply labor throughout the distribution.[28] For the childcare subsidy perturbation, the leak due to labor supply responses is significantly larger at 0.47. Most of this, 0.43, can be attributed to above-median income households. This number is higher than for the tax-transfer perturbation because in addition to the own-price effect on labor supply O_w, these households also work less due to a cross-price effect $X_{\mathcal{K}}$, see part 2(b) of Proposition 2: childcare becomes more expensive. The labor supply leakage among households below the median income is small but positive. Hence, the incentives to work more provided by the cheaper childcare price $X_{\mathcal{K}}$ are more than offset by the disincentive to work due to the steeper childcare price schedule O_w. The larger magnitude of the leaks for the above-median income household comes mostly from the fact that both incomes and tax rates are larger in this group, hence fiscal externalities are larger even when cross-price elasticities are of similar size (see Section 5.3).

The second row shows the magnitude of the leaks occurring as a result of households' changes in their demand for public childcare. For the tax-transfer change, these are the result of the cross-price effect of a change in the net wage, X_w: as households decrease their labor supply, they also decrease their use of nursery services – although some of the change in working hours is absorbed by changes in leisure. This is a saving for the government budget and partially mitigates the leaks occurring via the labor supply channel in the first row. Now looking at the right-hand side panel of Table 5 for the change to childcare subsidies, the -0.14 decrease in the MECR coming from the childcare adjustments of households above the median income represents the sum of the own-price effect of the price of childcare $O_{\mathcal{K}}$ (better-off households consume less childcare as it got more expensive) and of the cross-price effect of the net wage X_w (as the childcare fee schedule has become steeper, the net wage is lower at the margin). Both these adjustments create savings in terms of childcare subsidies and a reduction in the MECR. For households below the median, childcare has become cheaper and the own-price effect $O_{\mathcal{K}}$ constitutes a leak, which is only partly offset by the cross-price effect of the net wage, X_w, so that adjustments to childcare demand for this group add 0.09 to the MECR.

[28]Note that in contrast to the model in Section 2, income effects are at work. They mitigate the own-price effect on labor supply above the median and reinforce it below the median.

Comparing the contributions to the MECR coming from behavioral adjustments by households above and below the median income in the last row of Table 5, we observe that the bulk of the difference in MECRs between the two perturbations comes from adjustments by the better-off households: although they consume less public childcare, the fiscal externality caused by their decreased labor supply is large.

Table 6: Decomposition of labor supply effects into static and dynamic components

MECR Decomposition	**Tax-Transfer System**			**Childcare Subsidies**		
	$> y_{50}$	$\leq y_{50}$	$\sum$	$> y_{50}$	$\leq y_{50}$	$\sum$
Static labor supply	0.15	0.09	0.24	0.30	0.03	0.34
Dynamic labor supply	0.06	0.04	0.11	0.12	0.01	0.14
Total	0.21	0.13	0.34	0.43	0.05	0.47

Another angle we can take to better understand the sources of the leaks is to decompose the labor supply effects into static and dynamic components.[29] Table 6 summarises the results of the decomposition. We define as static the part of the leak that happens in the perturbation period: i.e., how do changes in labor supply in that period directly affect the leakage? Given the dynamic wage effects discussed above, these changes will also affect earnings in the future, both through wage effects and labor supply effects implied by those wage changes. This will affect the future budget and hence imply a leakage since our changes are budget neutral in the dynamic sense, i.e., in terms of the long-term government budget. Examining all the columns in Table 6, we see that the static effect is slightly more than twice as large as the dynamic effect.

6.2 Accounting for child development

We now incorporate effects of the mode of childcare on child outcomes into the MECR analysis as described in Section 2.5. While our data does not allow us to measure child outcomes, our model delivers counterfactual predictions on childcare demand responses to perturbations – and as discussed in Section 5.3, the model is broadly in line with existing quasi-experimental evidence in this regard. Relying on these counterfactual predictions, our procedure then mirrors the approach of Hendren and Sprung-Keyser (2020) to incorporate the estimated impact of the policy on long-run earnings of the children through child development.

First, to translate the changes in childcare demand into changes in child outcomes, we define the 'return' to public childcare by parental income, i.e., the effect of one year of public childcare on children's future earnings conditional on their parents' income.[30]

[29] We refrain from decomposing childcare responses into dynamic and static components. The reason is that the dynamic component is very small.

[30] The related step in Hendren and Sprung-Keyser (2020)'s analysis is described in their Appendix A.

As discussed above, these returns are not aligned with households' valuation of their children's nursery attendance, since households also value the time that the mother spends with their children for cultural or personal reasons. Empirical estimates of the impact of nursery attendance on lifetime income are scarce and we are not aware of any credible reduced-form evidence for Germany. We thus consider a range of possible combinations of returns. We simulate results for average returns for children growing up in above-median income households between -1% and +1% and average returns for their below-median income counterparts between 1% and 5%. These different returns align with the hypothesis that public childcare acts as an equaliser (Cornelissen et al. 2018). These ranges overlap with the returns estimated by Havnes and Mogstad (2011b, 2015), who used Norwegian data and a large public childcare expansion to obtain reduced-form estimates of around 3.5% and -0.5% average returns to childcare attendance for children in households with incomes respectively below and above the median household income.[31]

Second, to determine the child-development adjusted MECR, we translate the returns to one year of full-time childcare into increases in the net-present value of lifetime earnings. Within our setting, we assume that these returns apply to children attending public childcare. To incorporate this into the MECR analysis, we need to obtain the implied increases in lifetime earnings of children. Our approach shares large similarities with the projection method used by Hendren and Sprung-Keyser (2020) described in their Appendix I. We use recent estimates for Germany from Dodin et al. (2022) to obtain average child earnings as a function of the parental income rank when the parents were young. We augment these numbers with life-cycle profiles from Bönke, Corneo, and Lüthen (2015). This yields the life-cycle earnings profiles of children as a function of the parental income rank. Combining these with the returns illustrated in Online Appendix Figure F.7, we obtain the increase in the net-present value of lifetime earnings due to public childcare attendance. All steps are explained in detail in Appendix F.

Results. In line with Section 2.5, we now discuss two ways to incorporate child development into the MECR. First, we only consider the fiscal externalities imposed by child development, i.e., $\rho^c = 0$.[32] Panel A of Table 7 summarises the resulting MECRs of the tax-transfer system and of childcare subsidies for different combinations of returns to public childcare: for children in below-median income households (r_b^{child}) we use values of 1.0%, 3.0% and 5.0%, and for children in above-median income households (r_a^{child}): -1.0%, 0% and 1.0%. The results in Panel A of Table 7 emphasise that $MECR(\widehat{T}_{50}, 0)$ is almost unaffected, ranging between 0.27 and 0.30 compared to 0.28 without child development (see Table 5). The tax-transfer perturbation decreases childcare demand for all families.

[31]These numbers result from our own calculations based on the results in Havnes and Mogstad (2011b, 2015). Details of our derivation are shown in Appendix F and Figure F.7.

[32]We apply the same tax function as for parents. We assume that children are singles or single earners in a married couple. Results barely change if we assume that their spouse has a positive income.

Therefore, when returns to public childcare are positive (respectively negative), children's lifetime earnings are negatively (respectively positively) affected by the tax-transfer perturbation and cause negative (respectively positive) fiscal externalities. In Table 7 this translates into increases in $MECR(\widehat{T}_{50}, 0)$ when returns to public childcare increase. Only for the case of negative returns for children in above-median income families (r_a^{child} = -1.0% in the left column) the MECR is decreased from fiscal externalities of child development.

$MECR(\hat{s}_{50})$ of childcare subsidies changes more substantially when accounting for the fiscal externality of child development: it ranges from 0.31 to 0.45 compared to 0.42 without child development. The childcare subsidy perturbation decreases childcare demand for above-median income families and increases it for those below-median. Therefore, negative returns to public childcare for above-median children combined with positive returns for below-median children imply a (substantial) reduction in $MECR(\hat{s}_{50})$ due to the positive fiscal externalities implied by the higher life-time earnings of children along the entire income distribution. Overall, the tax-transfer system is a more efficient redistribution instrument at the margin if we consider only the fiscal externalities of child development (Panel A).

Second, as in Hendren and Sprung-Keyser (2020), we also account for the change in child utility through the change in the net-present value of after-tax lifetime earnings. We compute the augmented definitions of the MECR defined in (16) and (17), and consider the same Pareto weight on children as on parents, i.e., $\rho^c = 1$. Panel B in Table 7 summarises the results. Changes are qualitatively similar to those in Panel A, but are now amplified since direct effects on children's lifetime utility are considered as well as fiscal externalities. $MECR(\widehat{T}_{50}, 1)$ increases (for almost all combinations of returns) compared to $MECR(\widehat{T}_{50}, 0)$. This is driven by the decrease in childcare demand by below-median children who therefore suffer from lower future wages, which decreases their lifetime utility. Accounting for the compensating variations of children due to child development has a more marked impact on $MECR(\hat{s}_{50}, 1)$ since progressive childcare subsidies depress childcare demand for children in better off households and stimulates it for children in poorer households. When this is combined with much greater returns to childcare attendance for children in poorer households relative to returns for children in richer households, we see much smaller values of $MECR(\hat{s}_{50}, 1)$. For example, when $r_b^{child} = 5\%$ and $r_a^{child} = 0\%$, the $MECR$ of childcare subsidies is 0.10 (to be compared with a value of 0.42 when child development effects are ignored). Negative values of the MECR in Panel B imply that one Euro from above-median income household leads to redistribution of *more than* one Euro to those below-median.

Overall, when we account for compensating variations of children, the childcare subsidy perturbation is the more efficient redistribution instrument at the margin for most combinations of returns examined in Table 7. For $r_a^{child} = 0$ and $\rho^c = 1$, the return for below-median children required to equalise the two MECRs is 2.02%. We observe in this

Table 7: MECR with child development for: **tax-transfer system | childcare subsidies**

Panel A: $\rho^c = 0$			
r_b^{child} (in %)	r_a^{child} (in %)		
	-1.0	0.0	1.0
1.0	0.27 | 0.37	0.28 | 0.41	0.29 | 0.45
3.0	0.27 | 0.34	0.28 | 0.39	0.29 | 0.43
5.0	0.28 | 0.31	0.29 | 0.36	0.30 | 0.40
Panel B: $\rho^c = 1$			
r_b^{child} (in %)	r_a^{child} (in %)		
	-1.0	0.0	1.0
1.0	0.27 | 0.24	0.29 | 0.36	0.31 | 0.46
3.0	0.29 | 0.09	0.31 | 0.24	0.33 | 0.36
5.0	0.32 | -0.07	0.34 | 0.10	0.36 | 0.24
MECR without child development: 0.28 | 0.42			

Notes: Within each column, we show first the MECR of the tax-transfer system and second the MECR of childcare subsidies: $MECR^c(\widehat{T}_p, \rho^c) \mid MECR^c(\hat{s}_p, \rho^c)$ as defined in (16) and (17) for $p = 50$. Panel A includes the effects of the fiscal externality of child development, i.e., $\rho^c = 0$. Panel B includes the effects of the fiscal externality and of net income increases of child development, i.e., $\rho^c = 1$.

table that a greater difference in the returns to childcare attendance between children in below-median income households and children in above-median income households leads to a smaller gap between the MECR for childcare subsidies and the MECR for the tax-transfer system, and a negative gap when the difference in returns is large enough. Another way to illustrate these results is to show the sets of returns $(r_b^{child}, r_a^{child})$ for which the difference in MECRs is positive/negative. This is what we do in Figure 5 for $\rho^c = 0$ and $\rho^c = 1$.

As a last step, we compute the relative Pareto weight on children ρ^c for which the two policy instruments imply the same MECR, as defined in (16) and (17). Results are shown in Table 8. Again we observe that when returns to childcare attendance are much higher for below-median households than for above-median households, the condition on the Pareto weight on children for the childcare subsidy schedule to be the more efficient redistribution instrument is less stringent. For example, when returns are respectively 5% and -1% for below- and above-median households, the Pareto weight ρ^c on children only needs to be greater than 0.09 for $MECR(\hat{s}_{50}, \rho^c)$ to be smaller than $MECR(\widehat{T}_{50}, \rho^c)$.

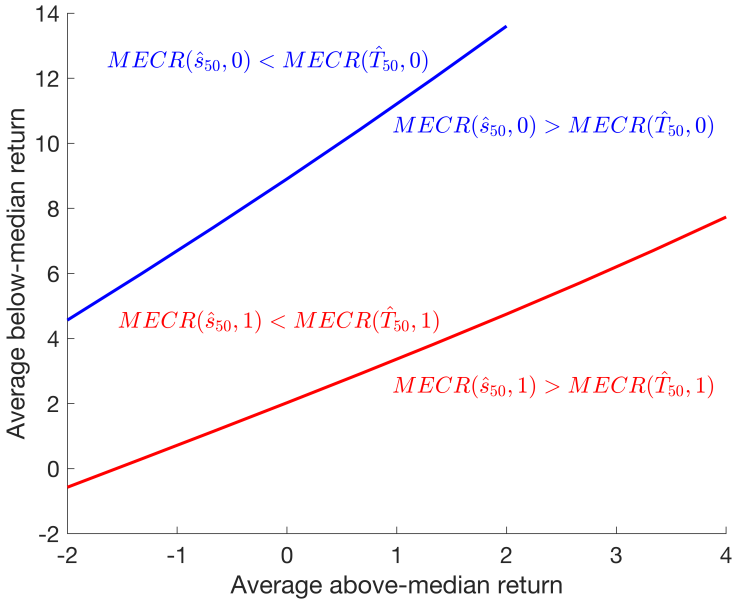


Figure 5: Order of MECRs for combinations of above- and below-median returns to public childcare

Whereas, when the pair $(r_b^{child}, r_a^{child})$ reflects a much smaller gap in returns at $(1\%, 0\%)$, the condition on the Pareto weight is $\rho^c \geq 2.33$ for progressive childcare subsidies to be more efficient at redistribution than the progressive tax-transfer schedule, at the margin.

Table 8: ρ^c to equalise MECR of childcare subsidy and the tax-transfer system

r_b^{child}	r_a^{child}		
	-1.0	0.0	1.0
1.0	0.82	2.33	11.81
3.0	0.27	0.58	1.23
5.0	0.09	0.23	0.47

6.3 Redistribution at other percentiles

We now use the flexibility of our approach and consider other threshold values for redistribution: we discuss here the MECR of redistributive perturbations around the 30th and the 70th percentile. Figure 6 illustrates our results for average returns above-median and below-median of 0% and 3% respectively, and $\rho^c = 1$, as an example. As before, we assume that returns decrease linearly in percentile – as in Havnes and Mogstad (2011b, 2015), see Online Appendix Figure F.7. The left panel of Figure 6 illustrates the case of $p = 30$ and the right panel the case of $p = 70$. Both can be directly compared to the case of $p = 50$ (seen above) in the middle panel of Figure 6. The patterns are very similar. If child development is not accounted for (blue bar), the MECR of the tax-transfer system is lower also for $p = 30$ and $p = 70$. The difference, however, is much larger for $p = 70$. If

child development is accounted for through fiscal externalities and through the impact on the compensating variation of the children, the ranking changes and the childcare subsidy turns out to be the more efficient instrument for redistribution at the margin for $p = 30$. For $p = 70$ however, accounting for child development decreases the gap between the two MECRs but the tax-transfer system remains the more efficient redistributive tool. This suggests that progressive childcare subsidies are only a more efficient means of redistribution when sufficiently targeted.

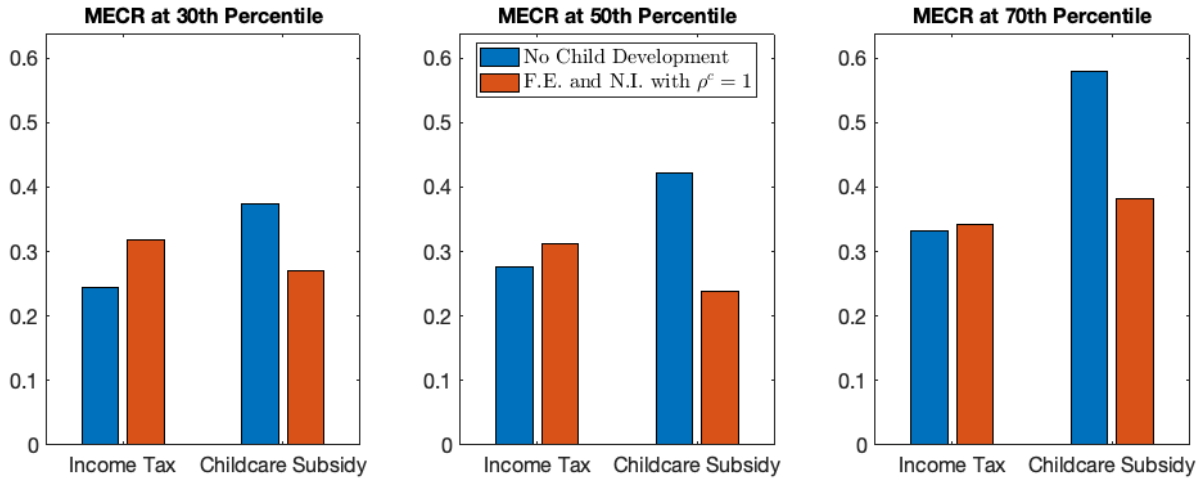


Figure 6: MECR for perturbations at different percentiles

Notes: The blue bars illustrate $MECR(\widehat{T}_p)$ and $MECR(\hat{s}_p)$ as defined in (8) and (9). The red bars illustrate $MECR(\widehat{T}_p, 1)$ and $MECR(\hat{s}_p, 1)$ as defined in (16) and (17). The left panel refers to $p = 30$, the middle panel to $p = 50$ and the right panel to $p = 70$. F.E. = fiscal externality, N.I. = net-income increases, both due to child development.

7 Conclusion

In this paper, we have incorporated the public finance approach of quantifying the efficiency costs of redistribution into a dynamic structural model of labor supply. Our measure of efficiency costs formalises the intuitive notion of Okun's leaky bucket (Okun 1975) in that behavioral adjustments in response to a perturbation cause some leaks in the redistribution process. We have compared the MECR of income-contingent childcare subsidies and of the tax-transfer system.

We have identified competing effects coming from maternal labor supply and child development. The maternal labor supply channel increases the MECR of childcare subsidies relative to the tax-transfer system. The child development channel decreases the MECR of childcare subsidies relative to the tax-transfer system. If one puts reasonably high Pareto weight on children and assumes returns to public childcare attendance that are in line with the quasi-experimental literature, childcare subsidies turn out to be the more efficient tool to redistribute from high-income to low-income families: for one Euro taken from

higher-income families, a greater amount reaches lower-income families.

Our decomposition of the efficiency costs makes it clear that – besides its immediate fiscal externalities – redistribution has a medium-term impact on the fiscal budget in the next 10-20 years via the impact of labor supply adjustments on mothers' future wages and a long-term impact on the income distribution of the children and the fiscal budget in 20-50 years' time via the impact of nursery demand adjustments on children's future earnings. Which of these effects is taken into consideration by the policymakers depends on the horizon over which they weigh the costs and benefits of potential reforms. And this will affect which tool is considered more efficient in achieving redistribution.

The structural model and the MECR play distinct roles in this analysis. The model is needed because the relevant objects are heterogeneous and dynamic: the MECR depends on income-specific own- and cross-price responses of labor supply and childcare demand, and on the fiscal consequences of maternal career effects. These objects cannot be recovered from a small set of aggregate elasticities. At the same time, the MECR remains the appropriate metric because our question is comparative and local: among redistributive margins already embedded in policy, which one loses less through behavioral and fiscal responses? Using the model to solve for the globally optimal childcare fee schedule would require additional normative choices and empirical extrapolations, in particular welfare weights and a structural technology for child development. We therefore view global optimal design as a distinct exercise rather than the object of this paper.

The general approach laid out in this paper could also be used to evaluate the dynamic efficiency costs of redistribution of other government policies such as the pension system, social housing, health insurance, etc. More conceptually, our paper also shows that one can harmonise equity concerns and social mobility concerns into one measure and therefore take a comprehensive perspective on policies that have distributional consequences on both the parent and the child generations.

References

Adda, J., C. Dustmann, and K. Stevens (2017). "The Career Costs of Children". *Journal of Political Economy* 125.2, 293–337.

Arnoud, A., F. Guvenen, and T. Kleineberg (2019). "Benchmarking global optimizers". Tech. rep. National Bureau of Economic Research.

Attanasio, O., H. Low, and V. Sánchez-Marcos (2008). "Explaining changes in female labor supply in a life-cycle model". *American Economic Review* 98.4, 1517–1552.

Authoring Group Educational Reporting (2018). *Education in Germany 2018*.

Bargain, O., K. Orsini, and A. Peichl (2014). "Comparing Labor Supply Elasticities in Europe and the United States New Results". *Journal of Human Resources* 49.3, 723–838.

Bastani, S., S. Blomquist, and L. Micheletto (2020). "Child Care Subsidies, Quality, and Optimal Income Taxation". *American Economic Journal: Economic Policy* 12.4, 1–37.

Bick, A. (2016). "The Quantitative Role of Child Care for Female Labor Force Participation and Fertility". *Journal of the European Economic Association* 14.3, 639–668.

Bick, A., B. Brüggemann, N. Fuchs-Schündeln, and H. Paule-Paludkiewicz (2019). "Long-term changes in married Couples' labor supply and taxes: Evidence from the US and Europe since the 1980s". *Journal of International Economics* 118, 44–62.

Bick, A. and N. Fuchs-Schündeln (2018). "Taxation and labour supply of married couples across countries: A macroeconomic analysis". *The Review of Economic Studies* 85.3, 1543–1576.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). "Female Labor Supply, Human Capital, and Welfare Reform". *Econometrica* 84.5, 1705–1753.

Boneva, T. and C. Rauh (2018). "Parental beliefs about returns to educational investmentsthe later the better?" *Journal of the European Economic Association* 16.6, 1669–1711.

Bönke, T., G. Corneo, and H. Lüthen (2015). "Lifetime earnings inequality in Germany". *Journal of Labor Economics* 33.1, 171–208.

Busse, A. and C. Gathmann (2020). "Free Daycare Policies, Family Choices and Child Development". *Journal of Economic Behavior & Organization* 179, 240–260.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children". *Journal of Political Economy* 123.2, 365–412.

Cesarini, D., E. Lindqvist, M. J. Notowidigdo, and R. Östling (2017). "The Effect of Wealth on Individual and Household Labor Supply". *The American Economic Review* 107 (12), 3917–3946.

Chetty, R. (2012). "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80.3, 969–1018.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins". *American Economic Review* 101.3, 471–475.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2018). "Who benefits from universal child care? Estimating marginal returns to early child care attendance". *Journal of Political Economy* 126.6, 2356–2409.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). "Interpreting the evidence on life cycle skill formation". *Handbook of the Economics of Education* 1, 697–812.

Cunha, F., I. Elo, and J. Culhane (2022). "Maternal subjective expectations about the technology of skill formation predict investments in children one year later". *Journal of Econometrics* 231.1, 3–32.

Dengler, K., B. Matthes, and W. Paulus (2014). "Occupational Tasks in the German Labour Market - An alternative measurement on the basis of an expert database". FDZ-Methodenreport 12/2014. Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB).

Dodin, M., S. Findeisen, L. Henkel, D. Sachs, and P. Schüle (2022). "Social Mobility in Germany". *Working Paper*.

Domeij, D. and P. Klein (2013). "Should Day Care be Subsidized?" *The Review of Economic Studies* 80.2, 568–595.

Eden, M. (2024). "The irrelevance of intergenerational altruism for social discounting". *CEPR Discussion Papers* 19147.

Ejrnæs, M. and A. Kunze (2013). "Work and wage dynamics around childbirth". *The Scandinavian Journal of Economics* 115.3, 856–877.

Elango, S., J. L. Garca, J. J. Heckman, and A. Hojman (2015). "Early childhood education". *Economics of means-tested transfer programs in the United States, volume 2*. University of Chicago Press, 235–297.

European Commission/EACEA/Eurydice (2019). *Key Data on Early Childhood Education and Care in Europe – 2019 Edition*. Eurydice Report. Luxembourg: Publications Office of the European Union. URL: https://eurydice.eacea.ec.europa.eu/sites/default/files/2023-01/EC0319375ENN.pdf.

Farhi, E. and I. Werning (2010). "Progressive Estate Taxation*". *The Quarterly Journal of Economics* 125.2, 635–673.

FDZ-SOEP - Forschungsdatenzentrum Sozioökonomisches Panel am Deutschen Institut für Wirtschaftsforschung (2019). *Sozio-oekonomisches Panel (SOEP), Daten der Jahre 1984-2017*. DOI: 10.5684/soep.v34.

FDZ-StABL - Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder (2020a). *Mikrozensus 2014, On-Site-Zugang*. DOI: 10.21242/12211.2014.00.00.1.1.1.

FDZ-StABL - Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder (2020b). *Mikrozensus 2018, On-Site-Zugang*. DOI: 10.21242/12211.2018.00.00.1.1.3.

Fricke, L. and D. Sachs (2024). "Efficiency Costs of Income Taxation: Income Effects and the Role of the Distribution". Tech. rep. Mimeo.

Gathmann, C. and B. Sass (2018). "Taxing Childcare: Effects on Childcare Choices, Family Labor Supply, and Children". *Journal of Labor Economics* 36.3, 665–709.

GESIS - Leibniz-Institut für Sozialwissenschaften (2017). *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 2016*. DOI: 10.4232/1.12796.

Golosov, M., M. Graber, M. Mogstad, and D. Novgorodsky (2021). "How Americans respond to idiosyncratic and exogenous changes in household wealth and unearned income". Tech. rep. National Bureau of Economic Research.

Golosov, M., A. Tsyvinski, and N. Werquin (2014). "A variational approach to the analysis of tax systems". Tech. rep. National Bureau of Economic Research.

Guner, N., R. Kaygusuz, and G. Ventura (2020). "Child-Related Transfers, Household Labour Supply, and Welfare". *The Review of Economic Studies* 87.5, 2290–2321.

Hank, K., K. Tillmann, and G. G. Wagner (2001). "Institutional child care in eastern Germany before and after Unification. A comparison with western Germany in the years 1990-1999". *Zeitschrift für Bevölkerungswissenschaft* 26.1, 55–65.

Havnes, T. and M. Mogstad (2011a). "Money for nothing? Universal child care and maternal employment". *Journal of Public Economics* 95.11-12, 1455–1465.

Havnes, T. and M. Mogstad (2011b). "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes". *American Economic Journal: Economic Policy* 3 (2), 97–129.

Havnes, T. and M. Mogstad (2015). "Is universal child care leveling the playing field?" *Journal of Public Economics* 127, 100–114.

Heathcote, J. and H. Tsujiyama (2021). "Practical optimal income taxation". Tech. rep.

Hendren, N. (2016). "The policy elasticity". *Tax Policy and the Economy* 30.1, 51–89.

Hendren, N. and B. Sprung-Keyser (2020). "A Unified Welfare Analysis of Government Policies". *The Quarterly Journal of Economics* 135.3, 1209–1318.

Hicks, J. R. (1939). "The Foundations of Welfare Economics". *The Economic Journal* 49.196, 696–712.

Ho, C. and N. Pavoni (2020). "Efficient Child Care Subsidies". *American Economic Review* 110.1, 162–199.

Mullins, J. (2022). "Designing cash transfers in the presence of childrens human capital formation". *Working Paper*.

OECD (2017). *Pensions at a Glance 2017*.

OECD (2020). "Is Childcare Affordable?" Policy Brief on Employment, Labour and Social Affairs. Paris: OECD. URL: https://web-archive-storage.oecd.org/aemint-web-archive-prod/web-archive/4b/4bda28da8febd677c0fb13407f2d2807e6593d725434aa282501e13pdf.

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Washington, DC: The Brookings Institution.

Piketty, T. (1997). "La redistribution fiscale face au chômage". *Revue française d'économie* 12 (1), 157–201.

Sachs, D., A. Tsyvinski, and N. Werquin (2020). "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium". *Econometrica* 88 (2), 469–493.

Saez, E. (2001). "Using Elasticities to Derive Optimal Income Tax Rates". *The Review of Economic Studies* 68 (1), 205–229.

Saez, E. and S. Stantcheva (2018). "A simpler theory of optimal capital taxation". *Journal of Public Economics* 162, 120–142.

Semykina, A. and J. M. Wooldridge (2010). "Estimating panel data models in the presence of endogeneity and selection". *Journal of Econometrics* 157.2, 375–380.

Statistisches Bundesamt (2012). *Finanzen der Kindertageseinrichtungen in freier Trägerschaft 2010*.

Stern, N. (2008). "The economics of climate change". *American Economic Review* 98.2, 1–37.

Turon, H. (2019). "Home production of childcare and labour supply decisions in a collective household model". IZA Discussion Paper 12148. IZA Institute of Labor Economics.

Virtanen, P., R. Gommers, T. E. Oliphant, et al. (2019). "SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python". *arXiv e-prints* 1907.10121.

Wales, D. J. and J. P. K. Doye (1997). "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms". *The Journal of Physical Chemistry A* 101.28, 5111–5116.

Wang, H. (2022). "Fertility and Family Leave Policies in Germany: Optimal Policy Design in a Dynamic Framework". *Working Paper*.

Online Appendix

A Proofs and Derivations of Section 2

In this appendix, we derive the fiscal externalities and the resulting formulas for $MECR(\widehat{T}_p)$ and $MECR(\hat{s}_p)$. Propositions 1 and 2 then directly follow. We start with the tax-transfer perturbation $\widehat{T}_p$ and then consider the subsidy perturbation $\hat{s}_p$.

A.1 Tax-transfer perturbation

We now consider all revenue effects of the tax-transfer perturbation $\widehat{T}_p$ in order to derive the budget neutral value for $\theta_b(\theta_a)$ First of all, the tax-transfer change has mechanical effects on government revenue given by:

$$\theta_a \int_{i:y_i>y_p} (y_i - y_p)\, di$$

and

$$\theta_b \int_{i:y_i\leq y_p} (y_i - y_p)\, di.$$

Fiscal externalities. In addition, it has budgetary effects through fiscal externalities. First of all, female labor supply for households above income y_p changes as follows due to increase of the marginal tax rate by θ_a:

$$\frac{\partial H_i}{\partial\left(1-\tau_i^H\right)}(-\theta_a) = -\frac{\partial H_i}{\partial w_{f,i}^{net}} w_{f,i}\theta_a = -\varepsilon^i_{H,w_f^{net}} \frac{y_{f,i}}{w_{f,i}^{net}}\theta_a.$$

Multiplying with the female wage and the labor wedges, yields the fiscal externality of this own-price effect on labor supply:

$$-\frac{\tau_i^H}{1-\tau_i^H}\varepsilon^i_{H,w_f^{net}} y_{f,i}\theta_a =: -O_w^i\theta_a. \tag{24}$$

Similarly, for households i with income below y_p the fiscal externality is:

$$O_w^i\theta_b = -\frac{\tau_i^H}{1-\tau_i^H}\varepsilon^i_{H,w_f^{net}} H_i\theta_b.$$

The change in the net-wage causes a change in childcare demand through a cross-price effect. Households i with $y_i > y_p$ change their childcare demand as follows:

$$-\frac{\partial N_i}{\partial w_{f,i}^{net}} w_{f,i}\theta_a = -\varepsilon^i_{N,w_f^{net}} \frac{N_i}{1-\tau_i^H}\theta_a.$$

The implied fiscal externality then reads as:

$$-\frac{\tau_i^N}{1-\tau_i^H} \cdot N_i \cdot \varepsilon^i_{N,w_f^{net}} \theta_a =: X_w^i \theta_a. \tag{25}$$

Similarly for households with $y_i < y_p$, it reads as

$$X_w^i \theta_b = -\frac{\tau_i^N}{1-\tau_i^H} \cdot N_i \cdot \varepsilon^i_{N,w_f^{net}} \theta_b.$$

Budget neutrality. To obtain $\theta_b(\theta_a)$, we add up all those fiscal effects and equate them to zero:

$$\begin{aligned}
&\theta_a \int_{i:y_i>y_p} (y_i - y_p)\, di + \theta_b \int_{i:y_i \le y_p} (y_i - y_p)\, di \\
&- \theta_a \int_{i:y_i>y_p} O_w^i di - \theta_b \int_{i:y_i \le y_p} O_w^i di \\
&+ \theta_a \int_{i:y_i>y_p} X_w^i di + \theta_b \int_{i:y_i \le y_p} X_w^i di = 0.
\end{aligned}$$

This implies

$$\theta_b(\theta_a) = \theta_a \frac{\int_{i:y_i>y_p} (y_i - y_p)\, di - \int_{i:y_i>y_p} O_w^i di + \int_{i:y_i>y_p} X_w^i di}{\int_{i:y_i \le y_p} (y_p - y_i)\, di + \int_{i:y_i \le y_p} O_w^i di - \int_{i:y_i \le y_p} X_w^i di}.$$

Inserting this into (8) yields:

$$MECR(\widehat{T}_p) = 1 - \theta_a \frac{\int_{i:y_i>y_p} (y_i - y_p)\, di - \int_{i:y_i>y_p} O_w^i di + \int_{i:y_i>y_p} X_w^i di}{\int_{i:y_i \le y_p} (y_p - y_i)\, di + \int_{i:y_i \le y_p} O_w^i di - \int_{i:y_i \le y_p} X_w^i di} \times \frac{\int_{i:y_i \le y_p} [y_p - y_i]\, di}{\int_{i:y_i>y_p} [y_i - y_p]\, di}$$

and hence

$$MECR(\widehat{T}_p) = 1 - \frac{1 - \frac{\int_{i:y_i>y_p} O_w^i di - \theta_a \int_{i:y_i>y_p} X_w^i di}{\int_{i:y_i>y_p} (y_i - y_p) di}}{1 + \frac{\int_{i:y_i<y_p} O_w^i di - \int_{i:y_i<y_p} X_w^i di}{\int_{i:y_i \le y_p} (y_p - y_i) di}}. \tag{26}$$

This result implies the statements in Proposition 1.

A.2 Subsidy perturbation

We now consider all revenue effects of the change in subsidy $\hat{s}_p$ in order to derive the budget neutral value for $\sigma_b(\sigma_a)$. First of all, we have mechanical effects on government revenue given by:

$$\sigma_a \int_{i:y_i>y_p} (y_i - y_p)\, N_i di$$

and

$$\sigma_b \int_{i:y_i \le y_p} (y_i - y_p)\, N_i di.$$

Note that here the difference is the multiplication with N_i since the burden or relief implied by this perturbation depends on how much public childcare N_i the household demands. In addition, it has budgetary effects through fiscal externalities.

Fiscal externalities due to changes in net wage. First of all, female labor supply for households above income y_p changes as follows due to increase of the effective marginal tax rate by $\sigma_a N_i$:

$$\frac{\partial H_i}{\partial\left(1-\tau_i^H\right)}(-\sigma_a N_i) = -\frac{\partial H_i}{\partial w_{f,i}^{net}} w_{f,i}(-\sigma_a N_i) = -\varepsilon^i_{H,w_f^{net}} \frac{y_{f,i}}{1-\tau_i^H}(-\sigma_a N_i)$$

The fiscal externality is then given by

$$-O_w^i \sigma_a N_i = -\frac{\tau_i^H}{1-\tau_i^H} \varepsilon^i_{H,w_f^{net}} y_{f,i} \sigma_a N_i.$$

Similarly, for households with $y_i \leq y_p$, the fiscal externality is:

$$-O_w^i \sigma_b N_i = -\frac{\tau_i^H}{1-\tau_i^H} \varepsilon^i_{H,w_f^{net}} y_{f,i} \sigma_b N_i.$$

The change in the net-wage causes as change in childcare demand through a cross-price effect. Households i with $y_i > y_p$ change their childcare demand as follows:

$$-\frac{\partial N_i}{\partial w_{f,i}^{net}} w_{f,i} \sigma_a N_i = -\varepsilon^i_{N,w_f^{net}} \frac{N_i}{1-\tau_i^H} \sigma_a N_i.$$

The implied fiscal externality then reads as:

$$X_w^i \sigma_a N_i = -\frac{\tau_i^N}{1-\tau_i^H} \varepsilon^i_{N,w_f^{net}} N_i \sigma_a N_i.$$

Similarly for households with $y_i < y_p$, it reads as

$$X_w^i \sigma_b N_i = -\frac{\tau_i^N}{1-\tau_i^H} \varepsilon^i_{N,w_f^{net}} N_i \sigma_b N_i.$$

Fiscal externalities due to changes in net cost. The perturbation $\hat{s}_p$ increases net costs of households with $y_i > y_p$ by $\sigma_a\left(y_i - y_p\right)$. This as an own-price effect on childcare demand. Childcare demand of households with $y_i > y_p$ changes as follows:

$$\frac{\partial N_i}{\partial \mathcal{K}_i^{net}} \sigma_a\left(y_i - y_p\right) = \varepsilon^i_{N,\mathcal{K}_{net}} \frac{N_i}{\mathcal{K}_i^{net}} \sigma_a\left(y_i - y_p\right)$$

The resulting fiscal externality is given by:

$$-\frac{\tau_i^N}{\mathcal{K}_i^{net}} \varepsilon^i_{N,\mathcal{K}_{net}} N_i \sigma_a\left(y_i - y_p\right) =: O_{\mathcal{K}}^i \sigma_a\left(y_i - y_p\right). \tag{27}$$

For households with $y_i \leq y_p$, the net-cost *decrease* $\sigma_b (y_p - y_i)$by it is given by:

$$O^i_{\mathcal{K}} \sigma_b (y_i - y_p) = -\frac{\tau^N_i}{\mathcal{K}^{net}_i} \varepsilon^i_{N,\mathcal{K}_{net}} N_i \sigma_b (y_i - y_p).$$

Finally, we turn to related cross-price effects. Female supply of households with $y_i > y_p$ changes as follows

$$\frac{\partial H_i}{\partial \mathcal{K}^{net}_i} \sigma_a (y_i - y_p) = \varepsilon^i_{H,\mathcal{K}_{net}} \frac{H_i}{\mathcal{K}^{net}_i} \sigma_a (y_i - y_p)$$

This results in the following fiscal externality

$$\frac{\tau^H_i}{\mathcal{K}^{net}_i} \varepsilon^i_{H,\mathcal{K}_{net}} H_i \sigma_a (y_i - y_p) =: X^i_{\mathcal{K}} \sigma_a (y_i - y_p). \tag{28}$$

For households with $y_i \leq y_p$, it is given by:

$$X^i_{\mathcal{K}} \sigma_b (y_i - y_p) = \frac{\tau^H_i}{\mathcal{K}^{net}_i} \varepsilon^i_{H,\mathcal{K}_{net}} H_i \sigma_b (y_i - y_p).$$

Budget neutrality. To obtain $\sigma_b(\sigma_a)$, we add up all those fiscal effects and equate them to zero:

$$\begin{aligned}
&\sigma_a \int_{i:y_i>y_p} N_i (y_i - y_p)\, di + \sigma_b \int_{i:y_i \leq y_p} N_i (y_i - y_p)\, di \\
&- \sigma_a \int_{i:y_i>y_p} O^i_w N_i di - \sigma_b \int_{i:y_i \leq y_p} O^i_w N_i di \\
&- \sigma_a \int_{i:y_i>y_p} X^i_w N_i di - \sigma_b \int_{i:y_i \leq y_p} X^i_w N_i di \\
&+ \sigma_a \int_{i:y_i>y_p} O^i_{\mathcal{K}} (y_i - y_p)\, di + \sigma_b \int_{i:y_i \leq y_p} O^i_{\mathcal{K}} (y_i - y_p)\, di \\
&+ \sigma_a \int_{i:y_i>y_p} X^i_{\mathcal{K}} (y_i - y_p)\, di + \sigma_b \int_{i:y_i \leq y_p} X^i_{\mathcal{K}} (y_i - y_p)\, di = 0.
\end{aligned}$$

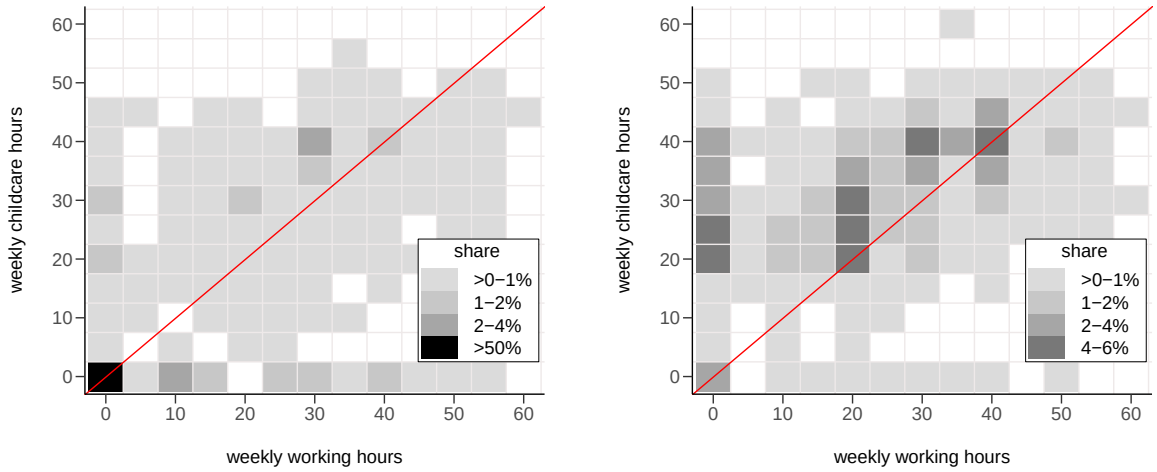
$$\sigma_b(\sigma_a) = \sigma_a \frac{\int_{i:y_i>y_p} N_i (y_i - y_p)\, di - \int_{i:y_i>y_p} (O^i_w + X^i_w) N_i di + \int_{i:y_i>y_p} (O^i_{\mathcal{K}} + X^i_{\mathcal{K}}) (y_i - y_p)\, di}{\int_{i:y_i \leq y_p} N_i (y_p - y_i)\, di + \int_{i:y_i \leq y_p} (O^i_w + X^i_w) N_i di - \int_{i:y_i \leq y_p} (O^i_{\mathcal{K}} + X^i_{\mathcal{K}}) (y_i - y_p)\, di}$$

$$\begin{aligned}
MECR(\hat{s}_p) =& 1 - \frac{\sigma^b(\sigma^a) \int_{i:y_i \leq y_p} N_i [y_p - y_i]\, di}{\sigma^a \int_{i:y_i>y_p} N_i [y_i - y_p]\, di} \\
=& 1 - \frac{1 - \frac{\int_{i:y_i>y_p} \left(O^i_w + X^i_w\right) N_i di - \int_{i:y_i>y_p} \left(O^i_{\mathcal{K}} + X^i_{\mathcal{K}}\right) (y_i - y_p) di}{\int_{i:y_i>y_p} N_i (y_i - y_p) di}}{1 + \frac{\int_{i:y_i \leq y_p} (O^i_w + X^i_w) N_i di - \int_{i:y_i \leq y_p} \left(O^i_{\mathcal{K}} + X^i_{\mathcal{K}}\right) (y_i - y_p) di}{\int_{i:y_i \leq y_p} N_i (y_p - y_i) di}}.
\end{aligned} \tag{29}$$

This result implies the statements in Proposition 2.

B Additional Stylised Facts

B.1 Childcare hours vs. working hours



(a) Children aged 0 – 2

(b) Children aged 3 – 5

Figure B.1: Maternal working hours vs. public childcare hours

Notes: Sample: mothers aged 20 to 65 that are not in education and live with a full-time working partner, conditional on having a child aged 0 – 2 for Figure B.1a or 3 – 5 for Figure B.1b. Source: 2009 to 2017 GSOEP, FDZ-SOEP (2019).

C Details on the Auxiliary Regressions

C.1 Childcare fees

C.1.1 Determinants of childcare fees

Child age. One of the important determinants of childcare fees in Germany is the age of the child. This is due to the fact that children of different ages visit different childcare institutions: from 0-2, they visit the nursery, and from 3-5 they visit kindergarten. Fees are usually higher for younger children since the costs of operating nurseries are higher than those of kindergartens.

Regional variation. Childcare fees in Germany differ further on a regional level because of two reasons: First, the fee schedules are set discretionary on a municipality level. Second, the German constitution ensures that the federal states bear the political responsibility for their respective education systems. Since public childcare is part of the education system, different federal states have implemented different regulations concerning the fee schedules.

Further determinants of childcare fees. Despite their autonomy, different states define in their legislation vastly similar determinants of childcare fees besides child age:[33]

1. *Household income:* In 11 out of 16 states the household income has to be used as a determinant and in two additional states it can be used.[34]

2. *Number of children in the household:* In 12 out of 16 states, childcare fees are determined conditional on the number of children in the household. Furthermore, in one additional state it can be used optionally as a determinant and two further states condition on the number of children in the household that attend nursery or kindergarten.

C.1.2 Estimation of the childcare fee schedule

We use data from the 2013, 2015, and 2017 GSOEP waves, which contains information on public childcare hours per day and monthly fees paid.[35] We normalise the monthly fees by the reported daily public childcare hours to extract the monthly fee of full-time public childcare, defined by an attendance of 8 hours per day or 40 hours per week. For this purpose, we assume linearity of childcare fees in hours.

[33] See Authoring Group Educational Reporting (2018): *Education in Germany 2018*, Section C2, p. 70–71 and Table C2-15web.

[34] Note that different municipalities differ in their definitions of household income: First, municipalities differ in using the net or gross household income. Second, they also condition on household income of different years (current year or previous years).

[35] In terms of the sample construction, this estimation is based on the same sample as laid out in Section 4.2.1 and D.1.

Given that we also observe a fraction of households paying zero fees, we estimate a Tobit model of childcare fees as a function of gross household income, which we also interact with the number of siblings.

We use the following linear model to estimate the childcare fee schedule reflected in the structural model by $p(j, K, y)$ separately for each child age bracket j:

$$\begin{aligned} p_{nt} = \; & \alpha + \beta_1 y_{nt} + \beta_2 \left(y_{nt} \times \mathbb{1}\Big\{\text{one sibling with age} < 17 \text{ in HH}\Big\}_{nt} \right) \\ & + \beta_3 \left(y_{nt} \times \mathbb{1}\Big\{\text{two siblings with age} < 17 \text{ in HH}\Big\}_{nt} \right) + \epsilon_{nt}. \end{aligned} \tag{30}$$

The dependent variable p_{nt} is the monthly fee that household n would pay for full-time childcare (40h/week) for a child aged j in year t. The interaction terms of gross household income with indicators for the number of siblings capture discounts granted to families with multiple children.[36] Our empirical model thereby closely reflects the current childcare fee schedule regulation as laid out in the previous section.

We estimate equation (30) as a Tobit regression with censoring at €0 and €725, the lowest and highest observed monthly childcare payments in our data.[37] We abstract from regional variation to keep the state space of the structural model tractable. But also if we extend the above regression (30) by state fixed effects and a dummy for living in an urban region to capture different levels of subsidies across regions, the resulting slopes of childcare subsidies with income remain unchanged compared to the following baseline results in Table C.1.

Results. The results of the Tobit regressions are summarised in Table C.1. Monthly childcare fees increase significantly in gross household income for all age brackets. Average fees are estimated to be highest for the youngest children, who require the most intensive care. The presence of siblings implies a significant reduction of the income gradient for $0-2$ and $3-5$ year olds, decreasing it by more than half if two siblings live in the household.

Figure C.2 shows the estimated fee schedule. Childcare fees are slightly increasing in household income (between 2% and 3% at the margin) and decrease with the number of siblings. Furthermore, fees are higher for younger children.

Finally, we take tax deductibility of childcare expenditures into account. We adjust the childcare fee schedule by the tax change implied by the childcare expenditures of the respective household income assuming full-time use of childcare.

[36] As we are mainly interested in predicting childcare fees, we only include covariates that are in line with the institutional setup described above. The stand-alone sibling dummies are not included as they do not add any explanatory power.

[37] We observe a number of households not paying any fees for a positive amount of childcare hours. Furthermore, we cap the fees to the maximum observed value to ensure that the rescaling to full-time equivalent fees does not yield unreasonably high values.

Table C.1: Tobit estimation of the childcare fee schedule

	Dependent variable: *Monthly fee for public childcare attendance of 40h/week*		
	child age		
	0 – 2	3 – 5	6 – 8
gross HH income	0.036	0.022	0.028
	(0.0037)	(0.0015)	(0.0043)
gross HH income × 1 sibling in HH	-0.012	-0.0065	-0.0050
	(0.0028)	(0.0012)	(0.0035)
gross HH income × 2 siblings in HH	-0.022	-0.010	-0.0052
	(0.0040)	(0.0015)	(0.0041)
constant	90.0	66.3	9.34
	(18.3)	(6.38)	(16.2)
N	362	1950	626

Notes: Sample: Children attending public childcare for whom childcare fees and childcare hours are observed. Tobit regressions with censoring at €0 and €725. All values are adjusted to 2017 price levels. Source: 2013, 2015, 2017 GSOEP, FDZ-SOEP (2019).

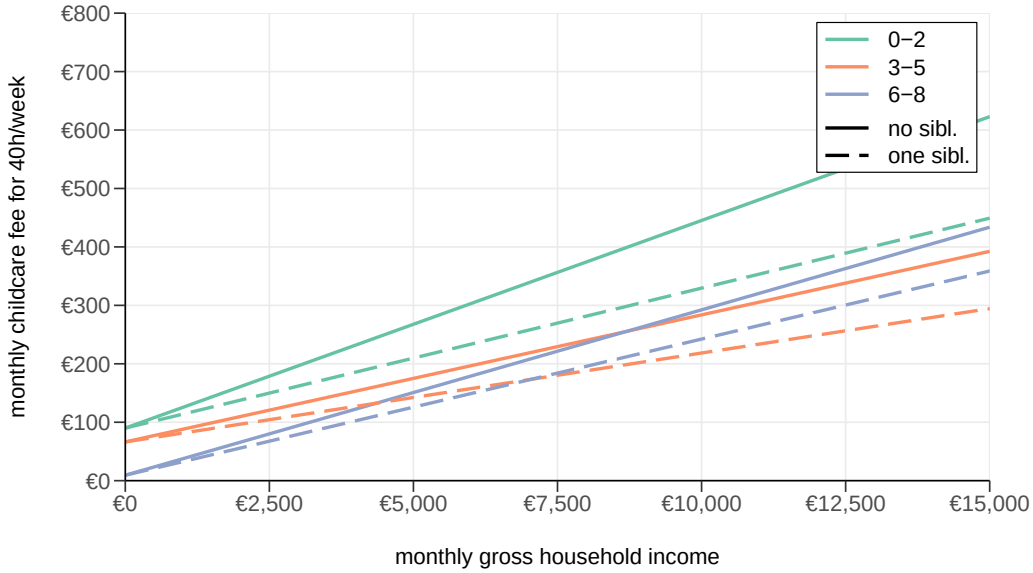


Figure C.2: Estimated childcare fee schedules

Notes: All values are adjusted to 2017 price levels. Source: 2013, 2015, 2017 GSOEP, FDZ-SOEP (2019).

C.2 Details on the fertility process

As introduced in Section 3.1, children are assumed to be born one at a time in any 3-year model period to mothers aged 20 to 40. We also restrict households to have at most three children. The determinants of fertility are the age and education of the mother and the number and ages of children already present in the family. The transition probability between family composition K and family composition K' faced by a household aged t, with an education level *educ* captures the (deterministic) ageing of existing children and the fertility hazard over the next period. Our estimate of the birth rate for this household is simply the sample average of birth events conditional on $(t, educ, K)$.

To make sure that we can identify also the less frequent fertility transition probabilities robustly, we compute them on an alternative larger data set, the German Microcensus. Specifically, we use the 2014 and 2018 Microcensus waves and focus on births taking place from 2012 to 2017.[38]

C.3 Details on the wage process

C.3.1 Potential wages for non-working females

For the imputation of potential wages of non-working females, we use the following static wage model:

$$log(w_{f,it}) = \mathbf{X}_{it}\boldsymbol{\beta} + u_{it}, \tag{31}$$

where $w_{f,it}$ is the wage of female i in period t and $\mathbf{X}$ contains the following Mincer-type covariates: linear and quadratic terms for age, full-time work experience, and part-time work experience. Furthermore, we include indicators for different education levels, namely an indicator for a lower track school degree and vocational training, an indicator for an A-level, and an indicator for a university degree. Additionally, $\mathbf{X}$ also contains the number of children below age 5, the overall number of children, an urban indicator, an indicator for living in former East Germany, and a full set of year indicators.

Wages are only observed if a woman works ($\text{participation}_{it} = 1$), which is determined by:

$$\mathbf{Z}_{it}\boldsymbol{\zeta} + \nu_{it} > 0, \tag{32}$$

where $\mathbf{Z}$ contains $\mathbf{X}$ along with a set of exclusion restrictions. Following Bargain, Orsini, and Peichl (2014) and in line with our model, we use as exclusion restrictions indicators

[38] We select the sample from the Microcensus data using the same restrictions as for our GSOEP survey data (see Section 4.2.1 and D.1). Sources: FDZ-StABL (2020a), FDZ-StABL (2020b). This yields us 71,165 observations of households aged 20 to 40.

for the presence of 0 – 2, 3 – 5, 6 – 8, 9 – 17, or 18+ year old children in the household. Furthermore, we include the husband's gross wage quintile and the net household income if the female chooses not to work.

In line with the selection correction procedure proposed by Semykina and Wooldridge (2010), we run a Probit version of equation (32) for each time period. In these, we also include the individual specific means of all covariates in $\mathbf{Z}$ across 2000 to 2017, denoted by $\overline{\mathbf{Z}}$:

$$Pr(\text{participation}_i = 1) = \Phi\left(\mathbf{Z}_i\boldsymbol{\zeta} + \overline{\mathbf{Z}}_i\boldsymbol{\xi}\right). \tag{33}$$

After estimating (33) for each year, we obtain the inverse Mills ratios λ_{it}, which we then use as control functions in the selection corrected version of the wage equation (31):

$$log(w_{f,it}) = \mathbf{X}_{it}\boldsymbol{\rho} + \overline{\mathbf{Z}}_i\boldsymbol{\xi} + \gamma\lambda_{it} + u_{it}. \tag{34}$$

With the estimated coefficients $\boldsymbol{\rho}$ and $\boldsymbol{\xi}$ at hand, we impute the wages of the non-working females.

C.3.2 Details on the wage process estimation

Using the 2000 to 2017 GSOEP data, we observe monthly gross labor income as well as contracted working hours.[39] This allows us to directly compute hourly wages for every individual that is working. For females who choose not to work, on the other hand, we do not observe any labor income and therefore, we impute their *potential* gross hourly wages using a selection-corrected wage model.

We then estimate the following equation for the wage process of females:

$$\begin{aligned} log(w_{f,it}) = \alpha + \beta_1 log(w_{f,it-1}) + \beta_2 \mathbb{1}\{lm_{it-1} = NP\} + \\ \beta_3 \mathbb{1}\{lm_{it-1} = PT\} + \beta_4 educ_i + \mathcal{A}(t) + \varepsilon_{it}^{w_f}, \end{aligned} \tag{35}$$

where $\mathbb{1}\{lm_{it-1} = NP\}$ and $\mathbb{1}\{lm_{it-1} = PT\}$ are dummy variables that indicate whether a woman i was either not working or working part-time in period $t-1$.

The coefficients β_2 and β_3 are of particular interest for our analysis since they measure the dynamic wage penalty from working less than full-time. β_4 captures the wage increase due to having obtained an A-level and $\mathcal{A}(t)$ is a third-order polynomial in age. Note that the implied wage process is a Markov process, where the individual wage is drawn from a log-normal distribution that depends on the previous wage, previous employment decision, age, and education. The estimated coefficients are shown in Table C.2.

[39] We extend the sample for the wage process estimation back until 2000 to ensure that we can robustly capture the key dynamics with a sufficient number of observations. Otherwise, we use exactly the same sample restrictions as described in Section 4.2.1 and D.1.

Table C.2: Estimation of female and male wage dynamics

	$log(w_{f,t})$	$log(w_{m,t})$
t	0.020	0.010
	(0.013)	(0.0093)
t^2	-0.0026	-0.0018
	(0.0017)	(0.0012)
t^3	0.000082	0.000059
	(0.000067)	(0.000048)
$higheduc$	0.076	0.027
	(0.0052)	(0.0036)
$\mathbb{1}\{lm_{i,t-1} = NP\}$	-0.18	
	(0.0065)	
$\mathbb{1}\{lm_{i,t-1} = PT\}$	-0.057	
	(0.0057)	
$log(w_{f,t-1})$	0.75	
	(0.0057)	
$log(w_{m,t-1})$		0.91
		(0.0040)
constant	0.70	0.30
	(0.032)	(0.023)
σ^2	0.28	0.20
	(0.0017)	(0.0012)

Notes: See Section 4.1.3 for the regression setup. 'educ' indicates having obtained at least an A-level, NP and PT denote not working and working part-time, respectively. Sample: mothers aged 20 to 65 that are not in education and live with a full-time working partner, non-participation wages imputed as described in Appendix C.3.1. Source: 2000 to 2017 GSOEP, FDZ-SOEP (2019).

D Details on the maximum likelihood estimation

D.1 MLE sample

Starting out with six waves of GSOEP data (2012 – 2017), we only keep households that are observed at least twice within this time frame. Then we allocate all children into the corresponding model child age brackets (see Section 3) and the household into the corresponding child-age structure K. Next, we assign each observation to a 3-year-spanning model period, ensuring that these line up with the evolution of the child-age structure K across time. Finally, we average all household variables of interest within the assigned model periods and only keep households with complete information for two model periods.

Further, we condition on observing the following covariates for every female: hourly wage if working, hourly wage of the partner, education (A-level or not), religion (Catholic at age 20), state of residence, predominantly living in an urban or rural area, demanding occupation (having primarily worked in an occupation where at least one-third of the tasks can be classified as analytic non-routine). Finally, we drop households that are either in the top 1% or bottom 1% of the male or female wage distribution to avoid distortions.

For the estimation, we operationalise the large state space as follows: to capture the age range from 20 to 80, we set up $t = 20$ 3-year model periods. Heterogeneity in male and female wages is captured by 5 and 11 grid points, respectively, education by 2 different levels, and the family structure K as introduced in Section 3.1 requires 18 state space points. The unobserved heterogeneity in g and α is captured by 20 grid points each, while 17 grid points are used for I reflecting a 2.5h grid $\{0, 2.5, 5, ..., 40\}$.

D.2 Identification

The identification is conditional on the calibrated and reduced form regression inputs (see Section 4.1), the homogeneous preference parameters (see Section 4.2.2), and the previously described assumptions for our maximum likelihood procedure. The distribution of $h = (g, I, \alpha)$ will be jointly and set identified.

There are three interacting ingredients that identify the time-invariant unobserved heterogeneity: i) cross-sectional variation in choices conditional on the same observed states, ii) the longitudinal dimension of our panel data, iii) using data not only from households with small children, but also from those with older children. The following paragraphs describe the three ingredients in more detail.

First, we observe households making different choices conditional on the same observed states s and constant characteristics x. Within our model, these differences in choices are therefore driven by differences in unobserved heterogeneity h. For illustration purposes, consider the example of a household with a single child aged 0 – 2 and a part-time working

mother that buys 20 hours of public childcare. From this household's choices in isolation, I is identified to be ≤ 20 hours, as otherwise the household would buy less public childcare. Nonetheless, I is only set identified: For a given preference for leisure α, the above choices could result from different combinations of g and I. A low preference for domestic childcare g relative to leisure α would be consistent with I close to 20 hours, i.e., the mother consumes leisure and does not provide much domestic childcare. On the contrary, a high preference for domestic childcare g relative to leisure α would be consistent with I close to 0 hours, i.e., the mother spends a lot of time on domestic childcare and little on leisure.

Now, let us consider variation in the two choices which helps to identify the distribution of unobserved heterogeneity: i) A higher amount of public childcare bought implies a higher preference for leisure, lower preference for domestic childcare and decreases the upper limit of the amount of informal childcare. ii) A decrease in the amount of public childcare bought implies that the household's informal childcare use I is strictly positive because otherwise the household would be unable to cover the childcare need while the mother works part-time. iii) If the mother were to work full-time, that would imply a lower preference for leisure, a lower preference for domestic childcare, and would point-identify I at 20 hours. (iv) If the mother would be not working, that would reflect a higher preference for leisure without necessarily affecting g and I as the household still consumes 20 hours of public childcare.

Turning to the second ingredient, using panel data is crucial for two reasons: i) The longitudinal dimension of the data and the associated temporal variation strongly facilitates identification because it allows to disentangle temporary shocks from the time-invariant unobserved heterogeneity. ii) Changes in the family composition over time also affect which dimension of heterogeneity matters in which period: Consider a household in which a child below 9 is present in one period but not in the other, i.e., either a new child is born in the second period or the youngest child is between 6 and 8 in the first period. Then, the preference for leisure (α) helps to explain the choices in both periods, whereas the preferences for domestic childcare (g) and the availability of informal childcare (I) help to identify the choices while a child that requires childcare is present. In addition, deterministic changes in the family composition, i.e., when at least one child between 0 and 8 is present in both periods, also facilitate the joint identification of g, I, and α.

Third, the estimation sample also includes households who have children without childcare need (age 9 and above) in both periods. For these, the only unobserved heterogeneity that matters is the preference for leisure α, which explains the variation in their labor market choices conditional on wages and other observed characteristics. Hence, this group adds significantly to the identification of the distribution of α independent of g and I.

The combination of all three just described ingredients allows us to credibly identify the joint distribution of g, I, and α.

D.3 Optimisation routine

To solve the optimisation problem numerically, we use the basin-hopping algorithm in combination with a Matlab built-in minimization routine for constrained target functions (*fmincon*). The basin-hopping algorithm is a stochastic global optimisation algorithm used in various fields (Chemistry, Applied Mathematics, ...), which was first introduced by Wales and Doye (1997).[40]

Intuitively, the procedure works as follows: We set an (arbitrary) initial starting point and solve for a (possibly local) minimum given the specified constraints on the parameters using *fmincon*. As we do not know the shape of the multidimensional objective function, we cannot be sure to have found the global minimum. To increase the likelihood of finding the global minimum, the basin-hopping algorithm then applies a random perturbation to the parameters of the previously found (potentially local) minimum and restarts the minimization routine *fmincon* at the perturbed parameters. The basin-hopping algorithm then compares the new minimum to the previous one and records the point with the lowest target function value as a candidate for the global minimum. The algorithm repeats the procedure, always keeping track of the point that yielded the lowest target function value, until either a predetermined number of iterations has been completed or the global minimum candidate did not change for a predetermined number of iterations. Trading off runtime and precision, we set the number of iterations in the basin-hopping algorithm to 1,000.

We verified our global optimisation routine by implementing the TikTak algorithm, a recent multistart global optimisation algorithm put forward by Fatih Guvenen and co-authors (see Arnoud, Guvenen, and Kleineberg (2019) for further information). Initially, the TikTak algorithm explores the parameter space uniformly (by setting so called Sobold points) and then, based on the gathered shape of the objective function, narrows its search to the most favorable areas. It then initiates local searches from specifically favorable points within the parameter space. The structure of our model allows us to set a high number of Sobol seed points (100,000), i.e. points in the multi-dimensional parameter space at which the function will be initially evaluated.

D.4 Results

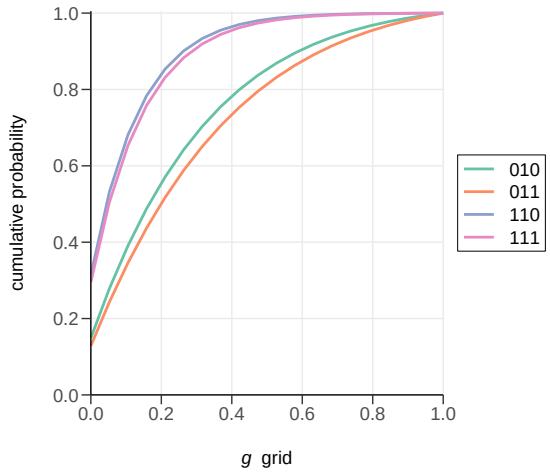
The estimated coefficients are shown in Table D.3.

[40] Our Matlab implementation of the basin-hopping algorithm follows the SciPy Python implementation (Virtanen, Gommers, Oliphant, et al. 2019).

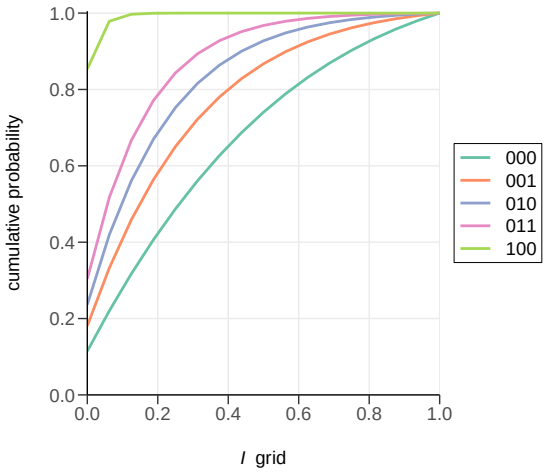
Table D.3: Maximum likelihood estimates

	domestic childcare (g)	avail. of informal childcare (I^{west})	avail. of informal childcare (I^{east})	leisure (α)
γ	-3.04	-1.27	-30.73	0.18
β_{east}	-4.37			
$\beta_{\text{demanding occup}}$	0.37			-0.80
β_{catholic}	0.54			
$\beta_{\text{high educ}}$		-2.80		-0.76
β_{urban}		-1.55		

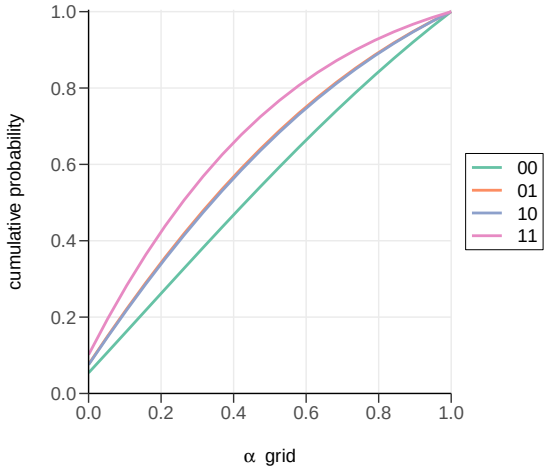
Notes: 'east' indicates having lived in former East Germany at some point; 'demanding occupation' indicates having primarily worked in an occupation where at least one-third of the tasks can be classified as analytic non-routine; 'catholic' indicates being Catholic at age 20; 'high education' indicates having obtained at least an A-level; 'urban' indicates living predominantly in an urban area. See Appendix D.3 for details on the optimisation and Appendix D.5 for an illustration of the sensitivity of the estimates. Note that we fix the variance of all normal distributions to 1.



(a) domestic childcare preference
covariates: east, skilled occupation, catholic



(b) informal childcare availability
covariates: east, education, urban



(c) leisure preference
covariates: high experience, education

Figure D.3: Cumulative distributions of unobserved heterogeneity

Notes: The legend of each subfigure indicates if the respective dummy – in the same order as the covariates listed below the subfigure – is 0 or 1. In case of the preferences for domestic childcare, we omit the plots for *no* skilled occupation because the implied difference by skilled occupation is small and would render the graph unreadable.

D.5 Sensitivity of MLE results

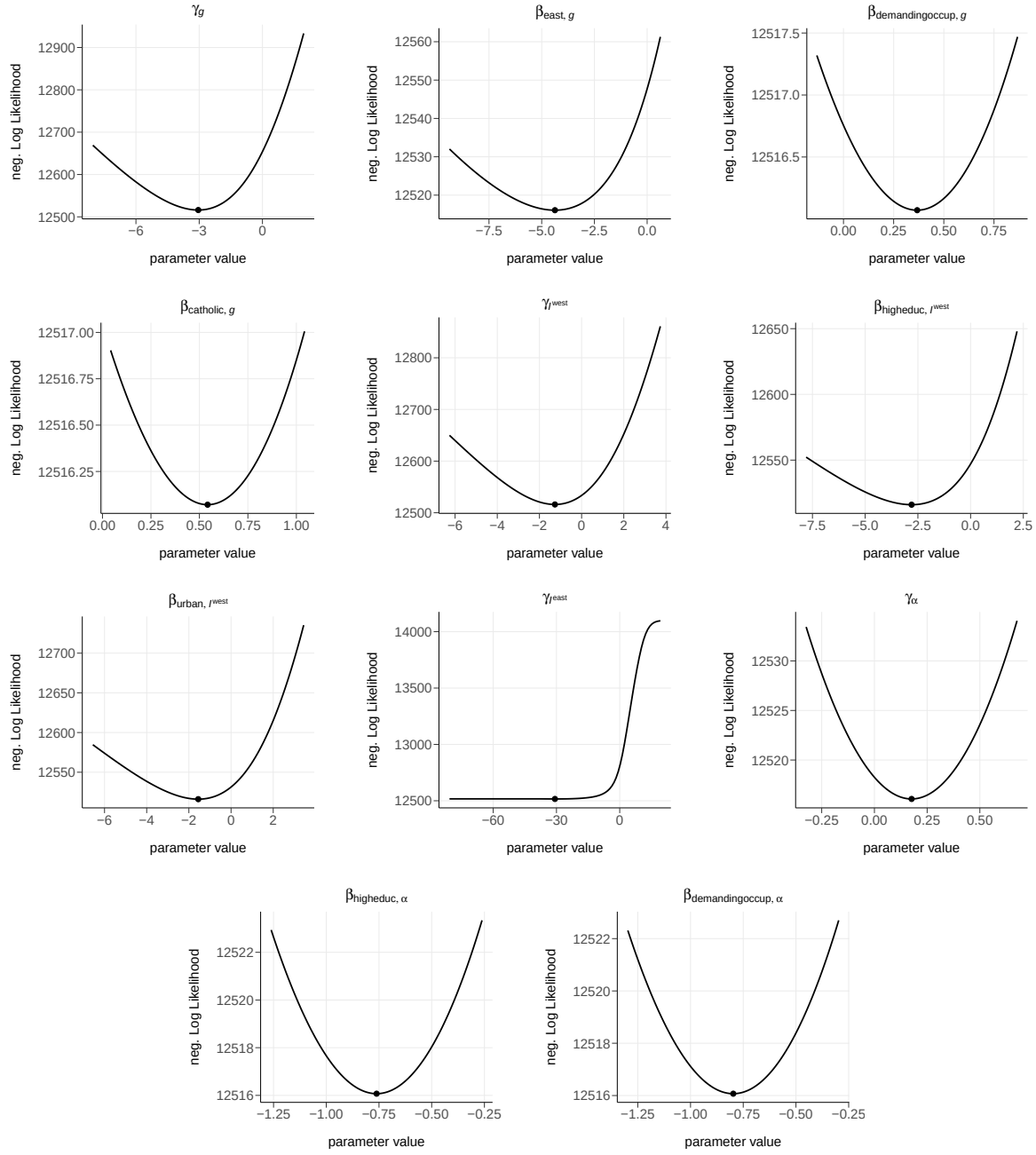


Figure D.4: Sensitivity of estimated structural parameters

Notes: Illustration of changes in the negative Log-Likelihood in response to small deviations of each parameter from its estimated value (labeled by black dot), keeping all other parameters at their point estimates. 'demanding occup' indicates having primarily worked in an occupation where at least one-third of the tasks can be classified as analytic non-routine. 'high educ' indicates having obtained at least an A-level. The intercept γ^I is not well identified. This likely has to do with the fact that in East Germany parents rely barely on grandparents for childcare. Note that the implied marginal distribution of I is effectively identical for other values of γ^I.

Object	Benchmark	Boot. mean	Bias	Boot. se	Δ_{95}	95% bound
Childcare MECR	0.4366	0.4387	0.0021	0.0216	0.0423	[0.3943, 0.4789]
Tax MECR	0.3126	0.3146	0.0020	0.0141	0.0276	[0.2850, 0.3402]
Childcare minus tax	0.1240	0.1241	0.0001	0.0157	0.0307	[0.0933, 0.1547]

Table D.4: Bootrap results

Notes: The table reports results from $B = 50$ paired household-level bootstrap replications. Each replication resamples complete two-period household histories with replacement, re-estimates the maximum-likelihood parameters governing the distribution of $h = (g, I, \alpha)$, and recomputes both dynamically budget-neutral perturbations at the median. "Bootstrap s.e." is the sample standard deviation across replications; for the final row, it is computed from the draw-by-draw paired difference between the childcare-subsidy and tax-transfer MECRs. "Bias" is the bootstrap mean minus the full-sample benchmark, and $\Delta_{95} = 1.96$ times the bootstrap standard error. The reported 95% intervals are normal-approximation intervals centered at the benchmark. To keep full re-estimation computationally feasible, the bootstrap uses 10 grid points for g and 5 for I, rather than 20 and 17 in the main specification. The benchmark values in this table therefore refer to the coarser-grid specification. Inputs calibrated or estimated outside the likelihood are held fixed.

D.6 Bootstrap inference.

In order to run the bootstrapped estimation in a reasonable time, we have adopted a coarser grid than in the main estimation: 5 grid points for I instead of 17 and 10 grid points for g instead of 20. The results are qualitatively similar and quantitatively close to the ones reported in the main text. We document here 50 true-MLE bootstrap draws. Each draw resamples households with replacement using the fixed draw-specific seed re-estimates the MLE parameters, and recomputes both budget-neutral Okun perturbations.

The results are summarized in Table D.4. The reported standard errors are the sample standard deviations across the 50 paired draws. The 95% bounds are normal-approximation bounds centered at the benchmark values, using $1.96 \times \widehat{\text{se}}$.

For the MECR difference, all 50 of 50 paired differences are positive. The bootstrap standard error of the benchmark difference is 0.0157; its centered normal 95% bound is [0.0933, 0.1547]. Thus, within this true-MLE bootstrap design, the evidence strongly supports a larger MECR for the childcare-subsidy perturbation than for the tax-transfer perturbation.

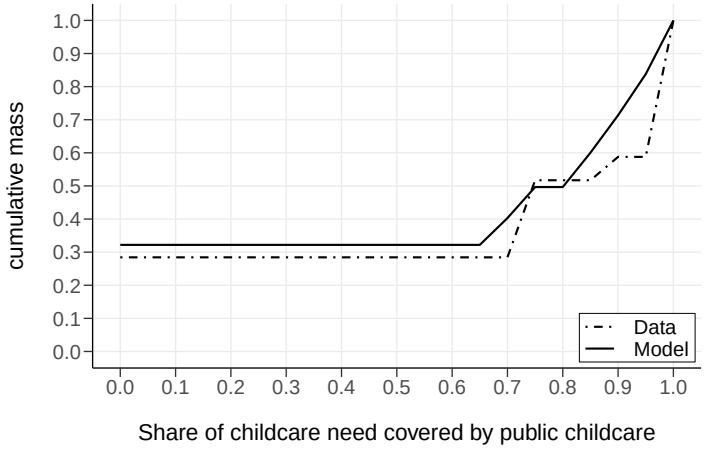
Estimated MLE Parameters Table D.5 reports the 14 estimated heterogeneity-distribution parameters. "Benchmark" is the full-sample MLE estimate; the remaining columns summarize the 50 independently re-estimated bootstrap MLEs.

E Additional model fit illustrations

Table D.5: Bootstrap variation in the maximum-likelihood estimates

Parameter	Benchmark	Bootstrap mean	Bootstrap s.e.
Preference for domestic childcare, g			
Intercept, γ^g	−9.4405	−9.3577	1.2953
East, β^g_{east}	−15.5038	−15.9267	3.4603
Demanding occupation, $\beta^g_{\text{demanding}}$	1.1047	1.2637	1.6371
Catholic, $\beta^g_{\text{catholic}}$	0.9799	0.9985	1.8226
Availability of informal childcare, I			
West intercept, $\gamma^{I,\text{west}}$	−3.4636	−3.3940	1.1998
High education, West, $\beta^{I,\text{west}}_{\text{high educ}}$	−7.1335	−7.5958	1.6486
Urban, West, $\beta^{I,\text{west}}_{\text{urban}}$	−4.5955	−4.6233	1.4894
East intercept, $\gamma^{I,\text{east}}$	−10.1052	−10.4194	1.5655
Preference for leisure, α			
Intercept, γ^α	−0.9976	−1.0879	0.5685
High education, $\beta^\alpha_{\text{high educ}}$	−3.1437	−3.1326	0.8702
Demanding occupation, $\beta^\alpha_{\text{demanding}}$	−3.2482	−3.1374	0.8402
Scale parameters			
σ_g	1.9999	1.9999	0.0001
σ_I	1.9999	1.9999	0.0001
σ_α	2.0000	2.0000	0.0000

Notes: The benchmark column reports the full-sample maximum-likelihood estimates under the coarser-grid specification used for the bootstrap. Bootstrap means and standard errors are computed across $B = 50$ household-level bootstrap replications, each of which resamples complete two-period household histories and re-estimates the MLE. Inputs calibrated or estimated outside the likelihood are held fixed.



(a) Youngest child in the family aged 6 – 8

Figure E.5: Model fit of families' public childcare demand

Notes: Sample as defined in Section 4.2.3. Data source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).

E.1 Model fit by household income

Table E.6: Model fit for labor supply by youngest child and household income

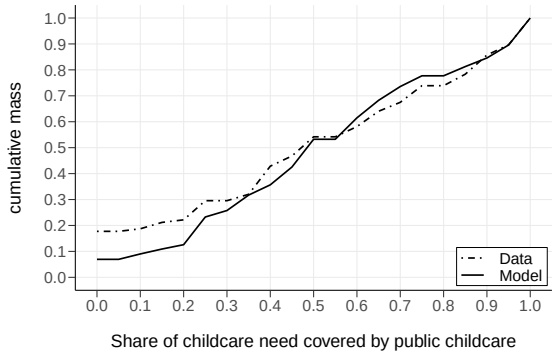
Panel A: Above-median income households

	Children 0 – 2			Children 3 – 5		
	NP	PT	FT	NP	PT	FT
Model	0.27	0.61	0.12	0.07	0.64	0.29
Data	0.34	0.56	0.11	0.07	0.71	0.22
	Children 6 – 8			Children 9+		
	NP	PT	FT	NP	PT	FT
Model	0.07	0.63	0.30	0.06	0.61	0.33
Data	0.07	0.70	0.23	0.07	0.54	0.39

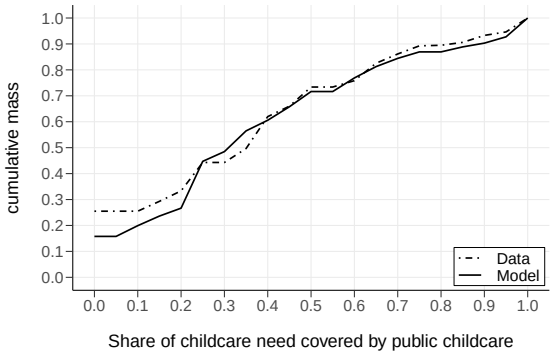
Panel B: Below-median income households

	Children 0 – 2			Children 3 – 5		
	NP	PT	FT	NP	PT	FT
Model	0.62	0.34	0.04	0.21	0.64	0.15
Data	0.65	0.32	0.03	0.26	0.59	0.15
	Children 6 – 8			Children 9+		
	NP	PT	FT	NP	PT	FT
Model	0.21	0.66	0.13	0.18	0.67	0.15
Data	0.21	0.65	0.14	0.25	0.53	0.22

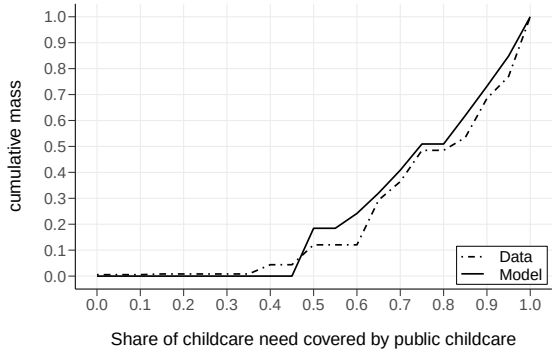
Notes: PT and FT denote the female working part-time and full-time, respectively. NP denotes non participation. Sample as defined in Section 4.2.3. Data source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).



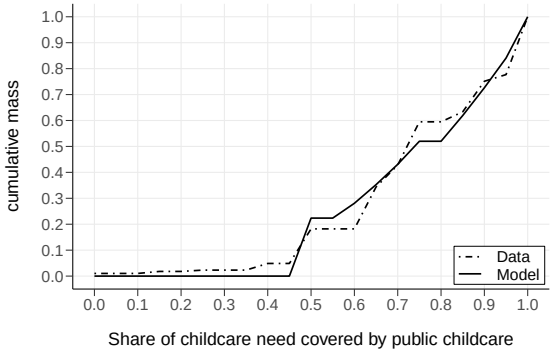
(a) Families with youngest child aged 0 – 2 and **above-median income**



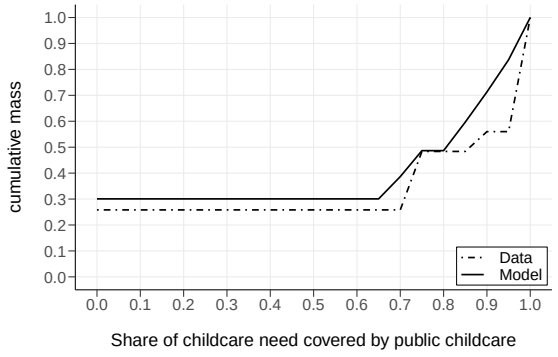
(b) Families with youngest child aged 0 – 2 and **below-median income**



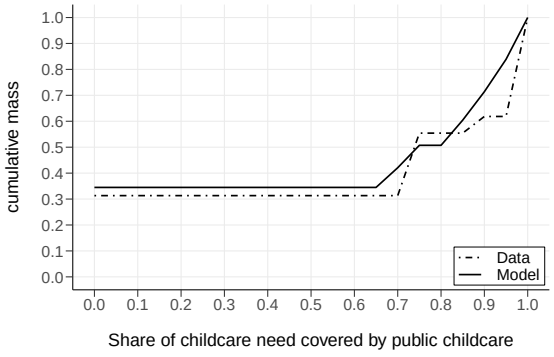
(c) Families with youngest child aged 3 – 5 and **above-median income**



(d) Families with youngest child aged 3 – 5 and **below-median income**



(e) Families with youngest child aged 6 – 8 and **above-median income**



(f) Families with youngest child aged 6 – 8 and **below-median income**

Figure E.6: Model fit of families' public childcare demand by household income

Notes: Sample as defined in Section 4.2.3. Data source: 2012 to 2017 GSOEP, FDZ-SOEP (2019).

F Child development analysis

F.1 Translating returns into net-present value increases

The next step consists of translating the returns to one year of full-time childcare into increases in the net-present value of lifetime earnings. Our approach shares large similarities with the projection method used by Hendren and Sprung-Keyser (2020) (see their Appendix I).

In a first step, we use recent estimates for Germany from Dodin et al. (2022) to obtain average child earnings as a function of the parental income rank. We take their rank-rank coefficients from Table 5 where child income is measured as individual labor earnings and parent income as gross family income.[41] This provides us with the estimated child rank at age 29-33 as a function of the parent rank. Based on the sample of the SOEP data used in this paper, we can then assign the child income in Euros that corresponds to these predicted child income ranks. In the final step, we extrapolate the lifecycle earnings profiles for each child earnings rank. We use the lifecycle profiles from Bönke, Corneo, and Lüthen (2015) estimated with German administrative pension data.[42] Assuming that children work from 25-60, this yields lifecycle earnings profiles from age 25-60 as a function of the parental income rank. Combining these with the returns illustrated in Online Appendix Figure F.7, we obtain the increase in the net-present value of lifetime earnings due to public childcare attendance.

F.2 Returns to childcare attendance

A large body of literature studies the effect of childcare on cognitive and non-cognitive skills and schooling outcomes of children.[43] However, evidence on long-term outcomes, particularly on labor income, is scarce. In this Appendix, we detail how we extrapolate the effect of attending one year of childcare on children's earnings during adulthood from reduced-form estimates of the impact of universal childcare on adult earnings in Norway by Havnes and Mogstad (2015).

Havnes and Mogstad (2015) exploit time and geographical variation in childcare provision in Norway induced by the Kindergarten Act in 1975. They obtain the effect of childcare attendance for children aged 3 to 6 on future adult earnings. They show how these returns vary with family income. Their regression estimates the reduced-form impact on all children from post-reform cohorts. Thus, their effects need to be interpreted as

[41] The intercept of this regression, which is not provided in this paper, equals 36.14.

[42] We assume the intermediate education level,high school + vocational training, for this extrapolation since Bönke, Corneo, and Lüthen (2015) estimate these profiles separately for high school only, high school + vocational training, and for college. Finally, we take the average of the male and female implied earnings growth rates over the lifecycle.

[43] See the surveys from Cunha et al. (2006) and Elango et al. (2015).

an intention-to-treat effect (ITT). In order to retrieve the impact of the treatment on the treated (TT), the ITT estimate is divided by the probability of treatment. Havnes and Mogstad (2015) define this as the percentage point difference between the increase in childcare coverage in the treatment and control municipalities in 1979, which equals 17.85%. Their estimate of TT thus suggests that each additional childcare slot induced by the policy reform increases adult earnings for children from low-income families on average by NOK 52′028.[44]

This number of 17.85% reflects a treatment effect per childcare slot offered but does not account for the gradual increase in childcare coverage over the years 1976 to 1979. Since we aim to quantify the yearly return of spending one year in public childcare, we need to adjust this estimate to reflect the actual increase in time spent in childcare caused by the reform. For example, a child born in 1976 enters childcare in 1979 and experiences the full treatment intensity of 17.85% during all three years of eligibility for public childcare. However, a child born in 1975 experiences a smaller treatment intensity during their first year of eligibility due to the gradual increase in childcare coverage. To account for this fact that the treatment intensity is hence not 3 years for all these children, we assume that the difference in childcare coverage between treatment and control municipalities equals 0.00% and then increases linearly until it reaches 17.85% in 1979; this is in line with Figure 3 on page 105 in Havnes and Mogstad (2015). The implied average treatment intensities (defined as an increase in the probability of attending childcare for 3 years) per birth year read as follows:

$$
\begin{aligned}
Prob_{1973}[Treat] &= \frac{0.00\% + 5.95\% + 11.90\%}{3} = 5.95\% \\
Prob_{1974}[Treat] &= \frac{5.95\% + 11.90\% + 17.85\%}{3} = 11.90\% \\
Prob_{1975}[Treat] &= \frac{11.90\% + 2 \cdot 17.85\%}{3} = 15.87\% \\
Prob_{1976}[Treat] &= \frac{3 \cdot 17.85\%}{3} = 17.85\%
\end{aligned}
$$

We further assume that equal numbers of children are born in each year and obtain an average treatment intensity of 12.89%. Based on the estimates by Havnes and Mogstad (2015),[45] we compute the TT of attending up to three years in childcare for income-poor and income-rich families. Taking a median family income to NOK 341′330.[46] and weighing estimates by population density, we obtain a TT for children born in income-poor families of NOK 41′068 and an effect of NOK −6′376 for children born in income-rich families.

[44] $NOK/EUR \approx 8$.

[45] See Figure 7 on page 110.

[46] Below-median families earn on average NOK 273′860, while above-median families earn NOK 440′420 ($NOK/EUR \approx 8$).

Subsequently, we express these treatment effects as returns to one year of public childcare by assuming constant compounded returns implying that $1 + r_{TT,1} = (1 + r_{TT,3})^{\frac{1}{3}}$. The return to one year in childcare for children born in families with below-median income equals 3.67%. For children from families with above-median income, the number equals to −0.52%. More generally, this approach yields the returns to one year of childcare as a function of parental income as illustrated in Figure F.7. We assume that these returns apply to one year of attending childcare for 40 hours per week and assume that these returns apply equally for child age between 0 and 5.

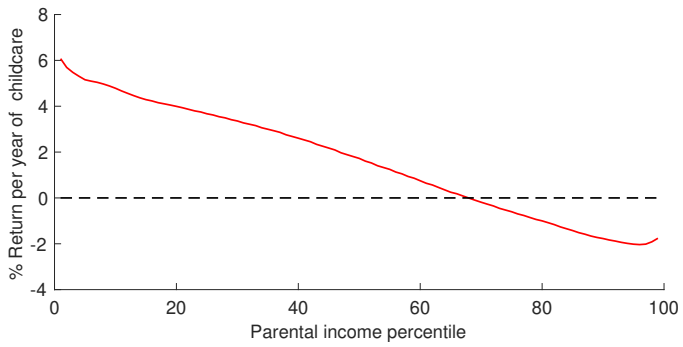


Figure F.7: Long-term returns to childcare attendance

Notes: This figure shows the return to children's earnings to one year of full-time public childcare attendance. The numbers are based on Havnes and Mogstad (2015).