Child Penalties in Time Use: A Global Perspective

Gabriela Deschamps* Amory Gethin† Camille Landais‡
Gabriel Leite-Mariante§

July 2026

Preliminary and incomplete: please do not circulate.

Abstract

Child penalties have been measured almost entirely through labour-market outcomes, yet paid work is under 15 per cent of adult waking time. Using harmonised time-use microdata from more than 80 countries, we estimate pseudo-event studies around first childbirth to trace how the full day reallocates across paid work, domestic work, childcare, and leisure. The first child adds roughly four hours of daily active time for the couple, financed out of leisure and sleep, and this cost persists fifteen years on. The reallocation falls overwhelmingly on mothers, and it is driven as much by domestic work as by childcare: mothers absorb almost all of the additional housework and most of the additional care, while fathers adjust little. This specialisation cannot be rationalised by wages alone — accounting for it requires implausibly large productivity differences favouring mothers — pointing to social norms rather than comparative advantage. In high-income countries, women's return to paid work coincides with falling domestic work rather than less childcare, and appliance diffusion alone accounts for about 30 per cent of that decline.

*UCL Stone Centre.
†World Bank.
‡London School of Economics.
§CUNEF.

1 Introduction

A recent literature shows that child penalties — the effects of parenthood on women relative to men — are a central driver of gender inequality in labour market outcomes, and in high-income countries account for most of it (Kleven et al., 2019, 2025). Child penalties have now been estimated for more than 130 countries and are a pervasive global regularity: men and women follow parallel trajectories until the birth of the first child, then diverge sharply and persistently. This literature has established parenthood as the central driver of gender inequality in the labour market. Yet it measures the impact of children almost entirely through a single lens: paid work.

Paid work accounts for less than 15 per cent of total waking time for adults aged 20–60 (Gethin and Saez, 2025). What parents do with the remaining 85 per cent is largely invisible to existing research. The child penalty in employment captures whether a mother reduces her hours in the labour market; it is silent on whether she absorbs additional domestic work, whether she loses sleep, whether the tasks she takes on are physically more demanding than those she displaces, and how all of this varies across societies at very different levels of development. Understanding the full time cost of children — how much additional work they create, who bears it, what kind it is, and how the answers differ across the world — requires data that no existing cross-country infrastructure provides.

This paper builds that infrastructure and uses it to study child penalties in the full allocation of time. In partnership with the Centre for Time Use Research at University College London, we construct a global harmonised time-use database covering more than 210 surveys across 80 countries, representative of approximately 90 per cent of the world population (Lamote de Grignon Perez et al., 2026). The database goes substantially beyond existing cross-country sources — such as the Multinational Time Use Study and the Harmonised European Time Use Surveys — by harmonising directly from raw country files, which allows a consistent and flexible set of activity categories across countries and over time. We apply the pseudo-event study methodology of Kleven (2022), adapted to time-use outcomes, to document what parenthood does to the full allocation of time around the birth of the first child. Three questions organise our analysis: what parental time investments do children require? How are those investments allocated within the household? And how do these patterns vary across countries and over the course of development?

We document five facts about parental time investment. First, the arrival of a first child is associated with around four additional hours of daily active time for the couple — a

20–25 per cent increase — financed entirely out of leisure including sleep. This cost is large and persistent: fifteen years after the first birth, the elevated time burden has not dissipated. Second, parental childcare intensifies sharply over the development path. Couple childcare time roughly triples between low- and high-income countries, a gradient that holds continuously from the world's poorest economies to the richest.

Third, childcare undergoes a qualitative transformation over development along three margins: how intensively parents care, what that care consists of, and who provides it. Care becomes more intensive relative to mere co-presence, as the share of time with a young child that mothers report as active childcare rather than passive supervision rises with income. It changes in content: in Tanzania, almost all parental childcare is physical care — feeding, washing, keeping the child safe; in the United States, play, educational support, and leisure activities account for the majority of childcare time. And it becomes more exclusively parental, as the broad alloparental network that characterised most of human evolutionary history gives way to an intensive nuclear-family model. Alloparents — siblings, grandparents, and other household members — provide around 30 per cent of all childcare at the bottom of the income distribution and under 10 per cent at the top, and the decline is driven almost entirely by young alloparents, whose contribution falls from roughly a quarter of all childcare to close to nothing while that of older household members is broadly flat.

Fourth, schooling and parental time investment are complements rather than substitutes. Institutional care expands enormously over development — weekly hours of schooling for children aged 0–14 roughly double between the poorest and richest countries — yet parental childcare expands alongside it rather than being displaced by it. Exploiting cross-country variation in the age structure of school access, we find that an additional hour per day of school access is associated with around 24 additional minutes of daily parental childcare for parents of children under five, with no detectable effect at primary or secondary level. Fifth, all of these time costs fall overwhelmingly on women. Women's paid work drops immediately and persistently, while their childcare and domestic work both rise steeply; fathers' profiles are barely disturbed, with paid work if anything rising modestly in some contexts, childcare increasing by less than an hour, and domestic work flat. Women thus bear almost the entire increase in domestic work and the majority of the increase in childcare — a complete reallocation of women's time that is the time-use foundation of the child penalties documented in the employment literature.

These five facts reveal a deep and persistent gender divergence in task allocation at childbirth, whose anatomy we dissect along four dimensions. Its most striking feature — one overlooked

by a child-penalty literature that works from employment data alone — is that unpaid domestic work matters as much as childcare in driving the divergence: food preparation, cleaning, and laundry all surge for women at childbirth and together contribute to the total gender gap in active time as much as childcare does, with food preparation alone accounting for roughly one additional hour per day. Second, the divergence is not only a matter of hours but of their character. Assigning each activity a physical-effort score following Cornwell et al. (2019), we find that women's average daily physical effort rises by around 20 per cent at first birth and stays persistently elevated while men's is essentially flat, so that standard child-penalty measures based on employment or earnings are silent on a welfare divergence that operates through the nature of work, not just its duration. Third, summarising each parent's day as an activity vector and measuring within-couple similarity, we find that children reduce similarity everywhere but by much less where the baseline is already low: mothers' and fathers' days are far more alike before parenthood in Europe than in Asia or Africa, so children generate *less* additional gender specialisation at low levels of development, because so much of it is already there. Fourth, the division of labour is deeply asymmetric. Studying within-household task complementarities in six countries spanning the income distribution, we find that when fathers work fewer paid hours they shift the freed time almost entirely to leisure, whereas mothers absorb changes across domestic work, childcare, and leisure at once; and fathers take on more paid work when mothers do more home production, but do not reciprocate when mothers move toward the market. Households do not rebalance symmetrically, and this rigidity holds at both ends of the income distribution.

This anatomy is not static: gender specialisation deepens and then unwinds over the course of development, and explaining that arc is a central aim of the paper. Placing mothers' and fathers' time allocations on the cross-country income distribution, we find that maternal paid work follows a U-shape — relatively high in the poorest economies, lowest at middle income, and recovering among the richest — while paternal paid work stays comparatively flat, so that specialisation first widens and only later narrows as countries grow. The recovery of maternal paid work at high income mirrors the decline of maternal *domestic* work almost exactly, whereas childcare rises monotonically for both parents and never falls. It is thus domestic work, not childcare, that recedes as women return to the labour market, and the two home-production margins behave so differently that any account of women's paid work over development must treat them separately.

To discipline these patterns, we set up a unitary household model of task allocation. The model yields three insights, each of which we take to independent data. First, the observed gender specialisation in childcare, with female-to-male ratios ranging from 2.5 to 7

across countries, cannot be rationalised by wage differences alone: the implied perceived-productivity ratio of mothers to fathers must exceed five to one, pointing to social norms rather than biological comparative advantage. Consistent with that interpretation, the cross-country gender gap in unpaid work is strongly correlated with directly elicited beliefs about maternal employment: a one percentage point increase in the share of respondents agreeing that a pre-school child suffers if the mother works is associated with a 3.3 minute wider daily gap. Second, rising parental investment in childcare over the development path requires the household's valuation of childcare to outpace rising wages, and when combined with strong norms, this force can widen rather than narrow gender specialisation, consistent with the first phase of the U-shape in maternal paid work. Our fourth fact reinforces this reading from a different direction: because institutional care complements rather than substitutes for parental time, the expansion of schooling over development does not relieve parents of care time. Third, the recovery of female labour force participation at high income requires a productivity revolution in domestic work, not cheaper market childcare. When parental and market childcare are complements, policies that expand outside care may have limited effects on maternal labour supply; the binding constraint is the domestic-work burden. We provide direct evidence for this channel by combining appliance-ownership and time-use information within the same surveys: holding appliance ownership at the rates observed in low-income countries would undo around 30 per cent of the observed decline in domestic work between lower-middle-income and high-income countries.

Our paper contributes to four literatures. It contributes most directly to the literature on child penalties, which has established the effects of parenthood on labour market outcomes at global scale (Kleven et al., 2019; Kleven, 2022; Kleven et al., 2025) and has begun to document related time-use patterns for a small number of countries (Aguilar-Gomez et al., 2026; Cuevas-Ruiz et al., 2026). That literature measures the impact of children almost entirely through paid work. Relative to it, our contribution is twofold. We open the full 24-hour day at global scale, which allows us to show that the labour-market penalty is one component of a much larger reallocation of time; and we use the granularity of diary data to micro-found the mechanisms the employment literature can only infer, separating childcare from domestic work, active care from supervision, and parental from alloparental provision. This is what lets us show that unpaid domestic work drives the gender divergence at childbirth as much as childcare does.

We contribute to the literature on parental investment and human capital formation, dating from Becker (1960) and developed by Doepke and Tertilt (2018); Doepke and Zilibotti (2019) to study how the nature of parenting varies over the course of development, and

connected both to the technology of skill formation (Cunha and Heckman, 2007) and to the strong gradient in parental time with children (Guryan et al., 2008). We provide the first global evidence on the quantity and nature of parental time investment over development. Our evidence on the transformation of childcare from physical care to educational investment, on the collapse of alloparental networks, and on the complementarity between schooling and parental time speaks directly to the distinction between survival-oriented and developmentally-oriented parenting at the heart of that literature; the complementarity result in particular suggests that the expansion of formal education raises rather than lowers the parental time price of a child.

We contribute to the literature on gender and time use over the course of development, which has used time-use data to study trends in leisure and home production (Aguiar and Hurst, 2007), to document the gender division of work across countries (Gottlieb et al., 2024), to trace the reallocation of time in Africa during structural transformation (Dinkelman and Ngai, 2022), and to measure the global evolution of working hours (Gethin and Saez, 2025). Relative to this largely cross-sectional and aggregate work, we bring the parenthood event study together with diary-level granularity — tasks, physical effort, co-presence, and alloparental provision — to a global sample, isolating how children reshape the gender allocation of time at every level of development.

Finally, we contribute to the literature on structural transformation and home production, which studies how household technology frees women for market work (Cowan, 1983; Greenwood et al., 2005) and how the marketisation of home production and the rise of services reshape gender gaps in work (Ngai and Petrongolo, 2017; Ngai et al., 2024). We measure home production and its gender allocation directly, on a common definition, across the full development range, and quantify how much of the decline in domestic work over development household technology can account for. Our findings suggest that the recovery of female labour supply at high income runs through a productivity revolution in domestic work rather than cheaper market childcare: because parental and market care are complements, the binding constraint is the domestic-work burden.

The paper proceeds as follows. Section 2 describes the data and the time-use infrastructure. Section 3 presents the pseudo-event study methodology. Section 4 documents five facts on parental time investment. Section 5 then turns to gender specialisation in three parts: its anatomy and well-being implications; how it evolves over the course of development; and a household task-allocation model with direct evidence on the two mechanisms it isolates. Section 6 concludes.

2 Data

2.1 Building a global map of time use

To document global patterns of time use around childbirth, we build a new global time-use database in partnership with the Centre for Time Use Research (CTUR-UCL). The database encompasses more than 210 surveys covering over 80 countries and is representative of approximately 90% of the world population. It goes substantially beyond existing cross-country harmonized time-use sources—such as the Multinational Time Use Study (MTUS) and the Harmonised European Time Use Surveys (HETUS)—by re-harmonizing data directly from individual country files, which allows for a more consistent and flexible set of activity categories across countries.

Our database draws on three types of sources. The first consists of existing cross-country harmonized time-use surveys, which we re-process from their underlying country files to impose our classification scheme. The second consists of individual country-level time diaries, obtained directly from national statistical offices. The third consists of household surveys with time-use modules, which provide recall-based information on time spent in broad activity categories. We treat diary-based data as the gold standard. In diary surveys, respondents record all activities over a full 24-hour period in short time intervals, producing a high-resolution record of the full daily time allocation. Recall-based surveys are included where diaries are unavailable, but are clearly distinguished from diary data in our coverage map.

Figure 1 presents a map of our data coverage by survey type. The map illustrates the breadth of our database across world regions. A striking feature is the extent of diary-based coverage—the gold standard for measuring time use—beyond the high-income world: diary surveys account for the majority of our country observations, and a substantial share of these are in low- and middle-income countries that have been largely absent from prior comparative work on time use.

[Figure 1]

Three features distinguish our database from prior harmonization efforts. First, it offers broader geographic and historical coverage than any existing source. Our database includes surveys from the 1960s onward (see Figure A2 in the appendix) and reaches countries that have been largely absent from comparative time-use research. Second, it is designed for versatility. Because we retain the full spell structure of the original diaries—rather than

collapsing activities into pre-specified totals, as existing harmonized sources do—researchers can recombine activity categories freely, extract spell-level information such as co-presence with children or time-of-day patterns, and construct novel attributes of each time unit. Third, the database is open source. In partnership with CTUR-UCL, we are releasing the harmonized microdata as part of the MTUS+ data release (Lamote de Grignon Perez et al., 2026), making the full infrastructure available to the broader research community.

2.2 Harmonisation

To harmonize time-use data across countries and survey formats, we convert all sources into a common long diary format. The unit of observation is an individual × time-slot × activity code: for diary surveys, each observation corresponds to a short interval and the activity recorded during that interval; for recall surveys, each observation corresponds to an activity episode with a reported duration. This common format allows us to apply a consistent classification scheme across all sources and is the input for all subsequent analyses.[1]

We classify activities using a three-level hierarchy, which is reported in full in Table A1 of the appendix. At the one-digit level, all time is allocated to one of three mutually exclusive categories: Active Time, which comprises both paid work and home production; Education; and Leisure, which includes basic personal care such as sleep and eating. At the two-digit level, Active Time is disaggregated into paid work, unpaid domestic work, unpaid care work, and voluntary work, while Leisure is broken down into leisure basics (sleep, eating and drinking, and personal care), entertainment, active leisure, and social activities. At the three-digit level, activities are classified at a finer grain: domestic work, for example, is separated into food preparation and clean-up, cleaning and upkeep, laundry and garment care, shopping and budget management, and production of goods for own use, while care work distinguishes dependent child care, dependent adult care, and care for non-dependent adults. Categories at every level are mutually exclusive and exhaustive: by construction, they sum to 1,440 minutes per day.

[1]Mapping the activity codes used by each national statistical office to our three-level hierarchy is a substantial undertaking given the diversity of classification systems across countries and survey waves. We address this using a large language model (LLM) classification pipeline that maps country-specific activity descriptions to our hierarchy in a consistent and scalable way. The pipeline runs each activity label through the model several times and assigns the modal label, which limits the influence of individual hallucinations; retrieval over the classification manual constrains the model to the admissible label set; and the output is reviewed and corrected manually for each country. After fine-tuning, agreement with human annotators is comparable to agreement between human annotators, with residual disagreement concentrated in intrinsically ambiguous activities. Appendix A.1 provides details.

2.3 Validation

We validate our measure of paid work hours against the global labour-force survey database constructed by Gethin and Saez (2025), which provides estimates of average hours worked separately for men and women across a large set of countries and years. For each country in our database, we identify the closest matching survey year in the Gethin–Saez database and compare average weekly hours of paid work by gender.

Figure 2 presents a scatterplot of this comparison. Our measure of paid work hours is on the horizontal axis and the Gethin–Saez labour-force survey estimate is on the vertical axis; men and women are shown in different colours, and a 45-degree line is included for reference. Across countries and genders, the observations cluster tightly around the 45-degree line, confirming that our harmonized database captures patterns of paid work consistent with those observed in dedicated labour-force surveys. The remaining deviations are driven by differences in survey timing and seasonality.[2]

[Figure 2]

3 Methods

Our global time-use database is cross-sectional: individuals are observed at a single point in time, either as parents or as childless adults, but never as both. The standard approach to estimating the effects of childbirth on outcomes—which traces individuals over time around the birth of the first child—requires panel data (Kleven et al., 2019). Panel time-use data are extremely rare even in high-income countries. To document global patterns of parental time allocation, we adopt the pseudo-event study approach of Kleven (2022), which was applied on a global scale to employment outcomes by Kleven et al. (2025). We provide a brief description here; a full treatment of the approach, including validation against true event-study estimates for employment outcomes, is in Kleven et al. (2025).

The idea is to use matching to convert cross-sectional data into a synthetic panel. For individuals observed with children, the age of the oldest child directly determines their position in positive event time ($t \geq 0$). The challenge is constructing the pre-birth counterfactual: these individuals are never observed without children. For each parent at event time $t = 0$ with age a and characteristics X_i, we match a childless individual observed at age $a - n$ with the same characteristics, providing a surrogate observation for event time $t = -n$.

[2]For instance, the survey timing in Tanzania in 2020 is likely to be affected by the Covid-19 pandemic.

Repeating this for $n = 1, \ldots, 5$ yields a pseudo-panel covering five years of pre-birth history. The matching variables are age, education (four categories), and marital status (four categories)—a parsimonious set that is consistently available across all of our data sources.

We estimate the following specification separately for each activity a and gender $g \in \{w, m\}$:

$$y^g_{ita} = \sum_{t \neq -2} \alpha^{a,g}_t \cdot \mathbf{1}[T_i = t] + \delta_{age} + \delta_{age \times c} + \delta_{year} + \delta_c + v^g_{ita}, \tag{1}$$

where y^g_{ita} is minutes per day spent on activity a by individual i of gender g at event time t, and T_i denotes years since the first birth. The equation is estimated separately for each gender and each activity. The event-time dummy at $t = -2$ is omitted, so the coefficients of interest, $\alpha^{a,g}_t$, measure the change in time use relative to two years before the first birth. δ_{age}, $\delta_{age \times c}$, δ_{year}, and δ_c denote age, age-by-country, survey-year, and country fixed effects, which together absorb life-cycle profiles, country-specific age patterns, and time trends.

For couple-level analyses of time investment in Section 4, we sum the event-study coefficients for men and women at each event time, $\hat{\alpha}^{a,m}_t + \hat{\alpha}^{a,w}_t$, to obtain the total household response. For the gender specialisation analyses in Section 5, we estimate equation (1) separately for men and women, which allows us to directly compare the event-time profiles.

4 Time Investment in Children

In this section we document five facts about the time costs of parenthood. We begin at the household level, establishing how large those costs are, how they vary across the development path, how the nature and provision of care change along it, and how care interacts with the expansion of formal schooling; we then show that within the couple, the costs fall very unequally on mothers and fathers.

Fact 1: Time investment in children is large

The first child creates a large and persistent increase in daily active time at the couple level. Panel (a) of Figure 3 presents the couple-level event study for total active time around first childbirth, pooling all countries in our database. At the moment the first child arrives, total active time increases by around four hours per day, a 20–25% increase over the pre-birth baseline. The effect is not transitory: the event study shows no meaningful convergence back toward pre-birth levels over the fifteen years that follow.

What accounts for this large increase in active time? Panel (b) of Figure 3 decomposes

this aggregate increase into its major components. Three patterns stand out clearly. The first is an immediate and sharp rise in childcare that peaks around birth and then declines gradually as the child ages, while remaining substantially above pre-birth levels throughout the observation window. The second is a sustained rise in unpaid domestic work that is less front-loaded than childcare but comparably persistent. Fifteen years after the first birth, unpaid domestic work is almost two hours per day above pre-birth levels. The two profiles in fact cross: from roughly eight years after birth onwards, domestic work rather than childcare is the largest single component of the additional active time a child generates. The third is a smaller but steadily growing increase in the production of goods for own use — subsistence agriculture, fetching water and fuel, and the manufacture of goods consumed within the household — which reaches around forty minutes per day per couple by the end of the window. This component is quantitatively concentrated in low-income settings, and it is invisible in labour-force data even though it is unambiguously work. Paid work, by contrast, shows little net change at the couple level. This last finding masks large heterogeneities between men and women, which we document fully in Section 5.

[Figure 3]

The additional four hours of active time per day come at the expense of leisure. Figure 4 presents the couple-level event studies for leisure and its components. Panel (a) shows an immediate and sustained decline in total couple leisure at birth. Panel (b) disaggregates leisure into its four two-digit components and documents that the decline is broad-based: basics (sleep, eating, personal care), entertainment, active leisure, and social activities all fall at birth and remain below pre-birth levels in the years that follow. Within basics, the decline is driven primarily by sleep, which falls by around one hour and remains persistently depressed (see Figure A3 in the appendix). The time cost of a first child is distributed across every component of rest and discretionary time.

[Figure 4]

Fact 2: Parental investment in childcare rises over development

The couple-level average documented in Fact 1 conceals enormous regional heterogeneity. Panel (a) of Figure 5 presents the childcare event study separately for six world regions. Sub-Saharan Africa shows the smallest response, with the couple total remaining below one hour per day. The Middle East and North Africa are somewhat higher but still modest. Asia, Europe, North America, and Latin America show much larger responses: in the first year after birth, couples in these regions devote 5–6 hours per day to childcare, around three

times more than in Sub-Saharan Africa and the MENA region.

Panel (b) of Figure 5 presents this gradient as a scatterplot across all countries in our database. The figure shows a strong and monotone positive relationship between GDP per capita and parental childcare investment that holds across all major world regions and is not driven by any particular country or population group.

[Figure 5]

Fact 3: The nature of childcare changes over development

The quantitative increase documented in Fact 2 is accompanied by a qualitative transformation in what childcare consists of. That transformation operates along three margins: how intensively parents care for the child during the time they spend with them, what the content of that care is, and who provides it.

We begin with intensity. Time spent *with* a child and time spent *caring for* a child are not the same thing. A parent may be co-present while cooking, farming, or resting, and diaries that record the presence of household members allow us to separate the two. For the subset of surveys that link co-presence to specific roster members, we compute, for mothers whose oldest child is aged seven or below, the ratio of time reported as childcare to total awake time spent in the presence of the child. Figure 6 plots this share against GDP per capita. The relationship is positive: the fitted slope implies that the share of time with a young child that is reported as active childcare rises by around three percentage points per log point of income ($\beta = 0.03$, $r = 0.33$), from roughly 29% at the bottom of the income range to roughly 37% at the top. Senegal and Mongolia sit at the low end and the Nordic and Central European countries at the high end. The relationship is imprecisely estimated — only thirteen surveys record co-presence at the level of detail required, and Bangladesh is a clear high outlier — so we regard this as suggestive rather than decisive. But its sign matters for interpreting Fact 2: part of the threefold rise in measured childcare over development reflects a genuine intensification of care during time with the child, not simply a reclassification of co-presence.

[Figure 6]

We turn next to content. We disaggregate childcare into four types: physical care, educational support, play and leisure, and other childcare. Figure 7 presents stacked event studies for three representative countries at sharply different levels of development.

Panel (a) presents Tanzania. At birth, childcare rises sharply and the increase is composed

almost entirely of physical care—feeding, bathing, and keeping the child safe. Physical care remains the dominant component throughout the post-birth window. Panel (b) presents India. Physical care remains large, but it is now joined by a growing share of play and some educational support, particularly as the child ages out of infancy. Panel (c) presents the United States, where the composition is reversed: play, educational support and other childcare together account for the majority of the childcare increase, and physical care, though large, is no longer dominant.

[Figure 7]

This progression reflects the distinction between survival-oriented and developmentally-oriented parental investment identified by Doepke and Tertilt (2018). At low income levels, the primary function of parental time is physical survival; as income rises, parents shift toward activities designed to build the child's human capital. The three-fold increase in childcare documented in Fact 2 therefore understates the full transformation in the parent–child relationship over the development path.

We turn finally to provision. Childcare in low-income settings is shared across a much broader network of household members than the parental couple, and this sharing largely disappears as income rises. Cooperative breeding—the systematic sharing of childcare beyond the biological mother—is a defining feature of human reproduction. Anthropological evidence on contemporary hunter-gatherer societies shows how pervasive this arrangement is: among the Aka of Central Africa, infants have on average 14 different caregivers besides their mother, and by 18 weeks they spend more time held by alloparents than by their mothers (Hrdy, 2024; Rilling, 2024).

We measure the contribution of alloparental care by constructing total childcare time per day among all household members aged 15 and above, and computing the share performed by carers other than the parents. We split alloparents into two groups: young alloparents, comprising siblings and other non-parent household members of child-bearing age or below, and old alloparents, comprising grandparents, extended family members whose own children are grown, and older members of the household. Figure 8 plots these shares against GDP per capita.

Two features stand out. First, the total alloparental share collapses over development. At the bottom of the income range, around 30% of all childcare is provided by someone other than the parents; the share is flat or mildly increasing up to roughly $5,000 per capita and then falls steadily, reaching under 10% at the top of the range. Second, the decline is almost

entirely a decline in *young* alloparental care. The share performed by siblings and other younger household members falls from around a quarter of all childcare to roughly two per cent, whereas the share performed by older household members is broadly flat at around five per cent throughout. The composition therefore reverses: at low income the alloparental network is dominated by the child's siblings and other young household members, while at high income what little alloparental care remains is grandparental. The collapse of the wider care network is thus closely tied to the demographic and structural transformations that shrink households and keep older children in school rather than at home.

[Figure 8]

Taken together, the three margins describe a single transformation. The village-based, co-operative model of child-rearing that characterised most of human evolutionary history has been progressively replaced, at higher income levels, by an intensive nuclear family model in which the parental couple bears the full burden of childcare. Care becomes more active relative to co-presence, more oriented toward the child's development than to its survival, and more exclusively parental. Each margin amplifies the finding in Fact 2: advanced-economy parents face not only more childcare in total but a more demanding version of it, and a larger share of it falling on themselves rather than on a wider network of household members. The next fact shows that the one substitute for parental time that does expand over development — formal schooling — does not offset this at all.

Fact 4: Education and parental time investment are complements

Facts 2 and 3 establish that parental childcare rises steeply over development, becomes more intensive, and falls ever more exclusively on the parents themselves as the informal network that shares it disappears. There is, however, one care input that expands dramatically over the same range: formal schooling. If institutional care substituted for parental time, the expansion of schooling would mechanically relieve parents of care hours and would work against the gradients we document. We show that the opposite is true. Schooling and parental time investment are complements, and strongly so for the youngest children.

We first document the expansion of institutional care. For 103 countries we construct total weekly hours of schooling received by children aged 0 to 14, combining age-specific enrolment rates, the number of days of instruction per year, and instructional hours per day, so that the measure reflects exposure actually received rather than nominal availability. Figure 9 plots this measure against GDP per capita. Weekly schooling hours roughly double over the development range, from around seven hours per week in the poorest countries in our

sample — Sierra Leone, Rwanda, Timor-Leste, Uganda — to around sixteen in the richest. The gradient is steep and continuous, with substantial residual dispersion at every income level.

[Figure 9]

We then place the two inputs on the same axes. Because childcare is reported by the parent rather than by the child, comparing it to schooling requires converting parent-reported care into care received by the child, which involves correcting both for the fraction of the parent's care time during which a given child is present and for double-counting when both parents care simultaneously. We estimate the presence corrections on the one survey that links co-presence to specific roster members and apply them everywhere; Appendix A.3 gives the construction and shows that the gradients we report survive even under an adverse assumption that maximally flattens them. Figure 10 then plots, for four age groups, fitted hours per day of schooling and of received parental childcare against GDP per capita.

The two inputs rise together, and the joint expansion is largest for the youngest children. For children under three, total adult-directed time rises from a little over one hour per day at the bottom of the income range to more than five hours at the top, of which schooling accounts for at most a fifth: essentially all of the increase is parental. For pre-primary children aged three to five the total rises from roughly one hour to four and a half, with schooling and parental care each contributing. Only at primary and secondary level does schooling account for the majority of the level, and even there parental care does not fall: it is flat to mildly rising throughout. Nowhere in the figure does institutional care displace parental time.

[Figure 10]

Cross-country co-movement is only suggestive: income raises both inputs, so their joint expansion is consistent with complementarity but does not establish it. For a sharper test we exploit variation across countries in the *age structure* of school access. Countries at similar income levels differ substantially in how early they enrol children and for how many hours, and this generates variation in exposure that hits parents of children of one age group but not another. Let child_i^k indicate that parent i has a child in age group $k \in \{\text{under } 5, \text{primary}, \text{secondary}\}$ and let A_c^k denote average daily hours of school access at level k in country c. We estimate

$$\text{childcare}_i = \sum_k \Big(\beta_k \, \text{child}_i^k + \delta_k \, A_c^k + \gamma_k \, \text{child}_i^k \cdot A_c^k \Big) + X_i' \theta + \lambda \, \text{GDP}_c + \varepsilon_i, \tag{2}$$

where childcare$_i$ is minutes per day of parental childcare and X_i contains age, age squared, an ever-married indicator, number of children, and education fixed effects. The coefficients of interest are the interactions γ_k, which measure how much more parental childcare responds to school access at level k for parents who actually have a child at that level, relative to parents whose children are all at other levels. The level terms δ_k absorb any country-level correlate of school access common to all parents, and GDP per capita is controlled for directly. The estimation uses 26 countries with population-adjusted weights and standard errors clustered by country.

Figure 11 reports the estimated γ_k, for all parents and separately by gender. For parents of children under five, an additional hour per day of school access is associated with around 24 additional minutes per day of parental childcare, and the estimate is statistically distinguishable from zero. The effect is larger for mothers, at around 34 minutes, than for fathers, at around 20, though the maternal estimate is far less precise. For primary-school children the point estimate is small and positive with a confidence interval comfortably spanning zero; for secondary-school children it is negative, around -7 minutes, and also indistinguishable from zero. Complementarity is thus concentrated at the pre-primary stage and fades as children age, which is consistent with the mechanism one would expect: early enrolment draws parents into school-adjacent activities — drop-offs, structured play, reading, homework support — whereas by adolescence formal instruction and parental time are largely independent.

[Figure 11]

This fact matters for two reasons. Substantively, it means that the rise in parental childcare over development documented in Fact 2 cannot be read as parents filling a gap left by weak institutions: parental and institutional investment expand together, and the parental time price of a child rises even as formal schooling expands. For policy and for the model of Section 5.3, it implies that expanding the supply of institutional care is unlikely on its own to free up much parental — and therefore maternal — time. We return to this point in Insight 3, where the complementarity between parental and outside care is precisely what rules out cheaper market childcare as the driver of women's return to paid work at high income.

Fact 5: The time implications of parenthood are highly gendered

The household-level time costs documented in Facts 1–4 are distributed very unequally between men and women. Figure 12 presents event studies around first childbirth separately for men and women in the same couple, showing paid work, unpaid domestic work, and

childcare.

[Figure 12]

Panel (a) presents the results for men. Paid work rises modestly at birth, a pattern partly driven by high-fertility settings such as India where male labour market participation increases at the time of first birth. Childcare rises, but by less than an hour. Unpaid domestic work is flat or slightly declines. The arrival of a first child does not substantially disrupt the prior time-use profile of fathers.

Panel (b) presents the results for women, and the contrast is stark. At birth, women experience an immediate and large decrease in paid work, a large increase in childcare, and a large increase in unpaid domestic work. The four additional hours of daily active time at the couple level documented in Fact 1 fall almost entirely on mothers, through all three channels simultaneously: less paid work, more childcare, and more domestic work. This complete reallocation of women's time at first birth is the time-use foundation of the child penalties documented in the employment literature and motivates the detailed gender specialisation analysis in Section 5.

5 Gender Specialisation

In this section we characterise the divergence between mothers and fathers documented in Fact 5, in three parts. First, we lay out the *anatomy* of specialisation and its well-being implications: event studies by gender across the full hierarchy of time-use categories; a measure of physical effort showing that the divergence extends beyond the quantity of time to the burden of the work itself; a synthetic index of activity-portfolio similarity summarising how far mothers' and fathers' days diverge; and the within-household task complementarities that reveal how asymmetric the division of labour is. Second, we study how specialisation *evolves over the course of development*, comparing the time allocations of mothers and fathers across the cross-country income distribution. Third, we interpret these patterns through a unitary household model of task allocation and confront its implications with independent evidence.

5.1 Anatomy and well-being implications

5.1.1 Anatomy of specialisation

We begin at the highest level of aggregation. Figure 13 presents event studies for active time, education, and leisure separately by gender, pooling all countries. Panel (a) shows total active time. Men and women both increase active time at childbirth, but by different amounts: active time rises by around two hours per day for women and around one and a half hours for men. The pre-birth trends for both genders are flat, confirming that the divergence is driven by the arrival of the child rather than pre-existing differences in trajectories. The gap—around half an hour per day, sustained over the full post-birth window—is a first measure of the gendered burden of parenthood. Panel (b) shows education. Women reduce education time at childbirth while men's education is essentially flat. Panel (c) shows leisure. Leisure falls for both parents by 1–1.5 hours per day at birth and remains persistently below pre-birth levels. Figures A4, A5 and A6 in the appendix present the event studies by region.

[Figure 13]

Figure 14 decomposes active time into its main components. Three patterns stand out. Panel (a) shows paid work. Women experience an immediate and persistent decline at childbirth, while men's paid work rises modestly—a pattern partly driven by high-fertility settings such as India, where male labour market participation increases at the time of first birth. This pattern is consistent with the child penalty in employment documented by Kleven et al. (2025) across 134 countries: the gendered divergence in paid work at parenthood is a global regularity.

Panel (b) shows unpaid domestic work. At childbirth, domestic work surges for women and remains elevated throughout the observation window. For men, domestic work is flat or, if anything, slightly declines. Panel (c) shows childcare. Women provide around four times more childcare than men worldwide. This ratio is established at birth and remains stable throughout the post-birth window. A striking finding emerges from comparing Panels (b) and (c): domestic work is as quantitatively important as childcare in explaining the total divergence between mothers and fathers. Prior work, which relies on employment data, captures only the paid-work margin and therefore misses this equally large domestic-work margin entirely. Figures A7, A8 and A9 in the Appendix present these event studies by region.

[Figure 14]

Figure 15 decomposes domestic work into food preparation, cleaning, and laundry. The largest component is food preparation: women add around one hour per day at childbirth, and this increase persists for at least fifteen years. Cleaning and laundry both increase more modestly but persistently. Together, these three components account for the bulk of the domestic-work gap documented in Panel (b) of Figure 14, and all of them fall almost exclusively on women. For men, none of the three shows a meaningful response to childbirth.

[Figure 15]

Figure 16 presents stacked event studies that show the full reallocation of the 24-hour day. Panel (a) shows men. The reallocation is modest: time shifts primarily from leisure into childcare, with little change in paid work and essentially no change in domestic work. Fathers adjust along a single margin. Panel (b) shows women, and the picture is fundamentally different. At birth, women undergo a complete reallocation of their 24-hour day: paid work and leisure both compress sharply, while childcare and domestic work expand to fill the space created. This is not a marginal adjustment—it is a reorganisation of the entire activity portfolio, established at birth and persistent for at least fifteen years. The same stacked 24-hour design appears separately by country Figure A10 in the Appendix.

[Figure 16]

5.1.2 Well-being implications: the physical burden of specialisation

The event studies above measure time quantities. But not all activities impose the same physical burden. To capture this dimension, we assign each activity a physical-effort score following the framework of Cornwell et al. (2019) and track the evolution of average daily effort around first birth, separately for men and women.

Figure 17 presents the result. For women, average physical effort rises at childbirth and remains persistently elevated: the increase amounts to around 20% relative to the pre-birth baseline. For men, the physical-effort profile is essentially flat around the event. The activities that women take on after childbirth — intensive physical childcare, food preparation, cleaning, laundry — are not only more time-consuming than the activities they displace, they are also more physically demanding. Standard child-penalty measures based on employment or earnings therefore understate the full welfare divergence between mothers and fathers. Women's days become not just longer but harder, while men's do not.

[Figure 17]

5.1.3 A synthetic index of specialisation

Physical effort captures the intensity of individual activities. We now ask a different question: how much do the overall daily routines of mothers and fathers diverge? To answer this, we construct a single index of within-household similarity. For each parent we represent daily time use as a vector in which each component is the total minutes spent in a given activity, excluding sleep. Within each couple, we compute the cosine similarity between the mother's vector $\mathbf{v}_f$ and the father's vector $\mathbf{v}_m$:

$$s_{hh} = \frac{\mathbf{v}_f \cdot \mathbf{v}_m}{\|\mathbf{v}_f\| \, \|\mathbf{v}_m\|}. \tag{3}$$

A value of one implies that both parents allocate their time identically; a value of zero implies completely orthogonal activity portfolios. We track the evolution of this index around the birth of the first child.

Figure 18 presents this event study by world region, together with the level of similarity in the reference year $t = -2$. Three insights emerge.

First, the pre-birth level of similarity differs enormously across regions. Two years before the first birth, the average couple's activity portfolios have a cosine similarity of 0.67 in Europe and 0.58 in Latin America, but only 0.43 in Africa and 0.38 in Asia. Gender segregation in daily activity is thus already far advanced before children arrive, and its extent varies systematically: in the low-income regions of our sample, prospective mothers and fathers already lead substantially different days.

Second, similarity declines at childbirth in every region we study. Gender specialisation in time use is a universal response to parenthood, not a feature of particular institutional or cultural settings.

Third — and this is the finding that the baseline levels bring out — the size of that decline is inversely related to how much specialisation is already present. The decline is largest in Latin America and Europe, at around 0.12 and 0.11 respectively, and smallest in Africa, at under 0.03, with Asia intermediate at around 0.06. The ordering mirrors the childcare gradients of Section 4: the regions where parental childcare investment responds most strongly to birth are the regions where the couple's days diverge most. Children therefore generate less *additional* gender specialisation at low levels of development, not because parenthood is more egalitarian there, but because the division of labour is already so unequal before the first child arrives. Indeed, even after the largest observed decline, European couples' days remain more similar than Asian or African couples' days were before children. Gender segregation

in daily roles is not created by the arrival of children; it is amplified by it, and amplified most where the starting point is most equal.

[Figure 18]

5.1.4 An asymmetric division of labour

A final structural feature of specialisation concerns how households rebalance — or fail to rebalance — when one parent's paid work changes. We study two types of within-household time complementarities. The first is within-person: when a parent works more paid hours, how do her or his own domestic work, childcare, and leisure change? The second is cross-spouse: when one parent shifts time toward paid work, how does the other parent's time allocation respond?

We estimate these complementarities in six countries spanning the income distribution: Bangladesh and Kenya at low income; South Africa and India at middle income; Spain and the United Kingdom at high income. The sample in each country consists of couple households with a youngest child aged ten or below.

Let w_h^g denote daily hours of economic work for parent of gender $g \in \{f, m\}$ in household h, and let $x_{h,j}^g$ denote daily hours in activity $j \in \mathcal{J} = \{\text{domestic work, childcare, leisure}\}$. Let Z_h be a vector of controls comprising age, education, marital status, household size, number of children, and day-of-week dummies. We define residuals $\tilde{w}_h^g$ and $\tilde{x}_{h,j}^g$ as the residuals from separate OLS regressions of each variable on Z_h. The within-person complementarity for parent g is the slope from the following regression:

$$\tilde{x}_{h,j}^g = \alpha_j^g + \beta_j^g \, \tilde{w}_h^g + \varepsilon_{h,j}^g. \tag{4}$$

The coefficient β_j^g measures how parent g's own time in activity j co-moves with their own economic work, conditional on observables. The cross-spouse complementarity is the slope from:

$$\tilde{x}_{h,k}^m = \alpha_{jk} + \gamma_{jk} \, \tilde{x}_{h,j}^f + \eta_{h,jk}, \tag{5}$$

where j indexes the mother's activity and k indexes the father's activity. We focus on four (j, k) pairs: (i) mother's childcare → father's economic work; (ii) mother's domestic work → father's economic work; (iii) mother's economic work → father's childcare; and (iv) mother's economic work → father's domestic work. Pairs (i)–(ii) test whether fathers shift toward paid work when mothers absorb more home production; pairs (iii)–(iv) test the converse. Both equations are estimated by OLS on a 100-bin binscatter of the residualised variables.

Figure 19 presents within-person slope coefficients for each of the six countries and the cross-country average, separately for men and women. For men, the substitution is concentrated almost entirely in leisure: when fathers work fewer paid hours, they reallocate the time to leisure, with little or no change in domestic work or childcare. Fathers face a two-task problem—paid work or leisure—and changes in their paid-work hours do not propagate into home production. For women, the pattern is different. When mothers work fewer paid hours, time flows simultaneously into domestic work, childcare, and leisure. A change in paid work triggers a reallocation across the full task portfolio.

[Figure 19]

Figure 20 presents the cross-spouse slope coefficients. When mothers devote more time to home production—domestic work and childcare—fathers increase their paid work correspondingly. But the complementarity is not symmetric. When mothers increase their paid work hours, fathers' domestic work and childcare do not increase in response. The adjustments that households make when a mother shifts toward the labour market are not the mirror image of the adjustments made when she shifts toward home production.

This asymmetry has a concrete implication. When female labour supply rises, the home-production tasks that mothers previously performed do not redistribute to fathers. The within-person evidence shows why: fathers who gain non-market time allocate it to leisure, not to domestic work or childcare. The cross-spouse evidence shows the same from the other direction: fathers do not absorb more home production when mothers shift toward the market. The two findings together document a structural rigidity in the household division of labour that holds across all six countries and at both ends of the income distribution.

[Figure 20]

5.2 Evolving nature of specialisation over the course of development

We now ask how the time allocations of mothers and fathers covary with the level of economic development across countries. Figure 21 presents scatterplots of average daily time by gender against GDP per capita for paid work, domestic work, and childcare.

Panel (a) shows paid work. The gradient for fathers traces a slight inverse-U over the development path. For mothers, paid work traces a U-shape: relatively high at low levels of development, declining at middle income, and recovering at high income. The first phase of development is therefore associated with a growing divergence between maternal and

paternal paid work; the second phase with a partial convergence.

Panel (b) shows domestic work. For mothers, domestic work follows an inverse-U: it rises with income, peaks at middle income, and then declines sharply. For fathers, domestic work is slightly increasing throughout. The decline in maternal domestic work at high income mirrors almost exactly the recovery of maternal paid work in Panel (a).

Panel (c) shows childcare. Both parents increase childcare monotonically over the development path, consistent with Fact 2. Women provide more childcare than men at all income levels, and the ratio, while narrower at high income, remains large throughout. Unlike domestic work, childcare does not decline — parents in advanced economies do more of it, not less. In Section 5.3 we use a simple household task-allocation model to account for these patterns.

[Figure 21]

5.3 Model and implications for gender inequality

The empirical patterns documented in the preceding sections are consistent with a wide range of stories about household behaviour. A mother who reduces paid work at childbirth may be responding to comparative advantage, to social norms, to the price of outside care, or to some combination of all three. To discipline which forces are quantitatively necessary to rationalise the data, we interpret the facts through a unitary household model of task allocation. The model is deliberately simple: its purpose is not structural estimation but rather to clarify the intuitions that the data deliver and to make precise the conditions under which the observed patterns can arise.

It is useful to state in advance what the model does. Household specialisation — which we define as the father's supply of paid work relative to household public goods, measured against the mother's — is a function of four families of forces:

$$\text{HH specialisation} \;=\; f\big(\{\beta\},\; w_r,\; \{\alpha_r\},\; \{\theta\}\big).$$

The vector $\{\beta\}$ collects the weights the household places on the different private and public goods it consumes, and the arrival of children operates through the weight on childcare: this is the *rising value of child care* channel. The relative wage w_r captures perceived career returns for the mother relative to the father, and the vector $\{\alpha_r\}$ captures the relative productivity of maternal and paternal inputs in each home-production task; together these are the *comparative advantage* channels, and because $\{\alpha_r\}$ is a perceived rather than a

physical object it is also where gender stereotypes enter. Finally $\{\theta\}$ collects technology and the availability of market substitutes for home production. The three insights below identify which of these forces must be doing the work, and Sections 5.3.4 and 5.3.5 bring independent evidence to bear on the two that the model singles out: stereotyped perceived productivities, and technological change in domestic work.

5.3.1 Setup

Two spouses $k \in \{m, f\}$ allocate time $x_i^k \geq 0$ across four tasks $i \in \{c, d, \ell, w\}$: childcare, domestic work, leisure, and paid work. Each spouse faces a time constraint, and the household faces a budget constraint:

$$x_c^k + x_d^k + x_\ell^k + x_w^k = T \quad \forall\, k \tag{Time}$$

$$w^m x_w^m + w^f x_w^f = y + p_c m_c \tag{Budget}$$

where w^k is the exogenous wage of spouse k, y is household consumption, and m_c is market childcare purchased at price p_c.

Childcare and domestic work produce household public goods via CES aggregators:

$$Q_c = \left[\alpha_c^m (x_c^m)^{\rho_c} + \alpha_c^f (x_c^f)^{\rho_c} + \alpha_c^o (m_c)^{\rho_c}\right]^{1/\rho_c} \tag{6}$$

$$Q_d = \left[\alpha_d^m (x_d^m)^{\rho_d} + \alpha_d^f (x_d^f)^{\rho_d}\right]^{1/\rho_d} \tag{7}$$

The parameters $\alpha_i^k \in (0, 1)$ are spouse k's *perceived productivity weight* in task i. These weights capture a combination of real productivity differences, social norms, and stereotypes — a point we return to below. The parameter $\sigma_i = 1/(1-\rho_i)$ is the elasticity of substitution between spousal inputs in task i, and α_c^o captures the productivity of market-purchased childcare.

The household maximises a joint utility function:

$$U = W(Q_c, Q_d) + \beta_m v(\ell^m) + \beta_f v(\ell^f) + \beta_y z(y) \tag{8}$$

where $W(Q_c, Q_d) = [\beta_c Q_c^\psi + \beta_d Q_d^\psi]^{1/\psi}$ aggregates the two household public goods, β_c and β_d are preference weights on childcare and domestic output, and $v(\cdot)$ and $z(\cdot)$ are CRRA sub-utilities over leisure and consumption: $v(\ell^k) = (\ell^k)^{1-\gamma_k}/(1-\gamma_k)$ and $z(y) = y^{1-\eta}/(1-\eta)$.

We model the arrival of children as a discrete shift in the preference weight: $\beta_c : 0 \to \bar{\beta}_c$. When $\psi < 1$ — so that childcare output and domestic output are complements — this raises

the marginal utility of both Q_c and Q_d, consistent with the simultaneous rise in childcare *and* domestic work that we document in Fact 1.

5.3.2 Key results

At an interior optimum, the shadow prices of time are proportional to wages: $\lambda^m/\lambda^f = w^m/w^f$. This is a standard result in unitary household models and pins down how the opportunity cost of non-market time varies across spouses.

Taking the ratio of first-order conditions for parental and market childcare yields the substitution margin between the two inputs:

$$\frac{m_c}{x_c^k} = \left(\frac{\alpha_c^o}{\alpha_c^k} \cdot \frac{w^k}{p_c}\right)^{\sigma_c} \tag{9}$$

A rise in spouse k's wage relative to the price of market care shifts the household toward outside childcare, with elasticity σ_c. This expression captures the standard intuition that cheaper or more productive market care reduces parental time in childcare.

The most informative result for our purposes is the cross-spouse ratio of time inputs within each household-production task. Taking the ratio of first-order conditions across spouses, all terms common to both — the preference shifters β_c and β_d, the complementarity parameter ψ, and the level of market care m_c — cancel, leaving:

$$\frac{x_i^m}{x_i^f} = \left(\frac{\alpha_i^m}{\alpha_i^f} \cdot \frac{w^f}{w^m}\right)^{\sigma_i}, \quad i \in \{c, d\} \tag{10}$$

This expression has three implications. First, wages matter: a higher male-to-female wage ratio w^m/w^f reduces the father's relative time in task i, with elasticity σ_i, because his opportunity cost of non-market time rises. Second, perceived productivities matter: a higher α_i^m/α_i^f raises the father's relative time in task i. Third, for childcare specifically, a decline in the price of market care p_c reduces *both* parents' time proportionally, leaving the within-task gender split unchanged. Crucially, the ratio in equation (10) is invariant to the preference shifters β_c and β_d and to the complementarity parameter ψ. These parameters affect the *levels* of childcare and domestic work but not the *within-task gender split*.

5.3.3 Three quantitative insights

We now show that, combined with the empirical facts, the model delivers three quantitative insights about the forces driving gender specialisation. Figure 21 presents the cross-country

gradients that Section 5.2 documents — scatterplots of mothers' and fathers' paid work, domestic work, and childcare against GDP per capita — and the three insights below state what norms, preferences over childcare, and productivity in domestic work must look like if the unitary model is to rationalise those panels. Each insight makes a statement about a parameter that is not directly observed; in the two subsections that follow we confront two of those statements with independent data.

Insight 1: Observed gender specialisation requires very strong norms

Panel (c) of Figure 21 shows that the female-to-male childcare ratio x_c^f/x_c^m ranges from 2.5 to 7 across countries. The gender wage ratio w^f/w^m is around 0.8 in low-income countries (International Labour Organization, 2024) and higher in more advanced economies, so setting $w^f/w^m = 0.8$ yields conservative lower-bound estimates of the required norm ratio. The elasticity of substitution σ_c is likely low: mothers and fathers perform *different types* of childcare — mothers concentrate on physical and routine care, while fathers are disproportionately involved in play and educational activities (Yeung et al., 2001; Lamb, ed, 2010) — suggesting strong complementarities rather than easy substitution, and hence $\sigma_c < 1$.

Given these parameter values, equation (10) implies that the only way to rationalise the observed childcare ratios is a perceived-productivity ratio α_c^f/α_c^m of at least 2 under the most conservative assumptions ($\sigma_c = 1$, $R^{\text{data}} = 2.5$), and substantially larger — exceeding 5 across most of the empirical range — when the inputs are complements ($\sigma_c < 1$), as the evidence suggests (see Appendix D.2 for details). A ratio of this magnitude is difficult to reconcile with biological comparative advantages alone, given that the underlying tasks — feeding, bathing, reading, playing — involve no inherent female advantage. The model therefore points to a central role for social norms and gender-stereotyped beliefs about who is more productive in childcare.

Insight 2: Critical role of β_c

The second insight addresses a tension in Figure 21. Panel (c) shows parental childcare rising monotonically in GDP per capita for both mothers and fathers, yet higher income also raises the opportunity cost of time at home. Holding β_c fixed, both spouses should substitute away from parental childcare as wages grow. The fact that they do the opposite requires the household's valuation of childcare to shift outward fast enough to dominate the opportunity-cost pressure. Appendix D.3 states the restriction formally: along the development path, β_c must grow faster than w^f (condition (A16)).

This rising β_c also helps explain Panel (a), where the paid-work gender gap widens over the first phase of development as mothers' market hours trace the downward leg of a U-shape. The key is that $R_c \equiv x_c^m / x_c^f$ is independent of β_c (equation (10)): an increase in β_c raises both spouses' childcare proportionally, leaving the within-couple split unchanged. Because the mother's baseline childcare is higher whenever $R_c < 1$, she absorbs a larger absolute increase and, through the time constraint, a larger loss of market time. The mechanism therefore requires three ingredients jointly: rising wages for both spouses, β_c outpacing w^f, and $R_c < 1$ — which, in the calibration, is driven primarily by the perceived-productivity asymmetry of Insight 1, with relative wages also entering through equation (10).

Insight 3: Closing the gender gap requires a domestic-work revolution

The most striking pattern at high levels of development is the large decline in mothers' domestic work time (Figure 21, Panel (b)). Childcare, by contrast, continues to rise throughout the development path for both parents. This means that outside childcare and parental childcare appear complementary rather than substitutable: cheaper market care cannot be the main driver of women's return to paid work. Fact 4 supplies direct evidence for this reading rather than leaving it as an inference from the aggregate gradients. Where children have more access to institutional care, their parents do not do less childcare — for children under five they do measurably more — so the low value of σ_c that the model requires is something we observe rather than assume.

Within the model, the only force that can rationalise the rise in female labour force participation at high income is a productivity revolution in domestic work. This could take the form of household appliances, market substitutes for cooking and cleaning, or changes in domestic infrastructure — all of which raise the effective productivity of domestic work and free up maternal time for the labour market. Critically, this insight is available only because our data separately identify domestic work and childcare. Conflating the two categories, as most labour-market analyses implicitly do, would obscure the mechanism entirely. Section 5.3.5 quantifies how much of the decline in domestic work over development one specific component of this revolution — the diffusion of household appliances — can account for.

5.3.4 Elicited gender norms and household specialisation

Insight 1 concludes that the observed allocation of childcare requires perceived-productivity ratios far in excess of anything a biological account could support, and attributes the residual to norms. That conclusion is an inference from the model: α_c^f / α_c^m is not observed, and any parameter recovered as a residual invites the objection that it is absorbing misspecification

rather than measuring beliefs. We therefore ask whether the implied norms line up with norms measured directly.

We match our time-use estimates to survey questions on gender roles, using as our measure the share of respondents in each country who agree that a pre-school child is likely to suffer if the mother works. Figure 22 plots the gender gap in unpaid work — mothers' minus fathers' daily minutes — against this share. The relationship is strong and precisely estimated: a one percentage point increase in the share agreeing is associated with a 3.3 minute wider daily gap (standard error 0.24). The fitted relationship spans the full range of the data, from gaps of under an hour per day in the Nordic countries, Canada, and the United States, where fewer than one respondent in five agrees, to gaps approaching five hours per day in Pakistan, India, Bangladesh, and Egypt, where agreement exceeds three quarters.

[Figure 22]

We do not read this as causal: norms and behaviour are jointly determined, and stated attitudes may in part rationalise an allocation arrived at for other reasons. The relevant point is narrower. If the large implied α_c^f/α_c^m of Insight 1 were an artefact of the functional form, there would be no reason for it to co-vary so tightly with independently elicited beliefs about maternal employment. That it does supports interpreting the perceived-productivity wedge as a norm rather than as a technology, and it also supplies the cross-country heterogeneity in $\{\alpha_r\}$ that Insight 2 requires: the parameter is not a constant of nature but varies systematically with what people in a country believe.

5.3.5 Quantifying the domestic-work revolution

Insight 3 concludes that the recovery of maternal paid work at high income must be financed by a productivity revolution in domestic work. We now put a number on one component of that revolution.

The pattern to be explained is in Figure 23. Average domestic work is around five hours per day in low-income countries, peaks at just over six hours in lower-middle-income countries, and then falls to a little over three hours in high-income countries. The inverse-U is the aggregate counterpart of the maternal gradient in Panel (b) of Figure 21, and the decline over its second half is the object of interest: three hours per day of home production disappear between the lower-middle-income and high-income groups.

Over the same range, ownership of household appliances rises steeply. Figure 24 reports ownership rates for refrigerators, clothes washers, microwaves, and dishwashers by income

group. Refrigerator ownership rises from around 44% in low-income countries to 99% in high-income countries, clothes washers from around a third to 96%, microwaves from 4% to 87%, and dishwashers from essentially zero to 65%. The transition is sharpest between the lower-middle-income and upper-middle-income groups, precisely where domestic work begins its decline.

[Figure 23]

[Figure 24]

To move from co-movement to accounting, we exploit the surveys that record appliance ownership and time use for the same households. Let S denote the set of appliances and let $D_{x,i}$ indicate that individual i's household owns appliance x. We estimate, at the individual level,

$$Y_i = \sum_{x \in S} \beta_x D_{x,i} + \sum_{\substack{x<y \\ x,y \in S}} \gamma_{xy} D_{x,i} D_{y,i} + \sum_{\substack{x<y<z \\ x,y,z \in S}} \delta_{xyz} D_{x,i} D_{y,i} D_{z,i} + \theta \, \text{inc_pctile}_i + \lambda_{\text{age}(i)} + \mu_{\text{country}(i)} + \nu_{\text{year}(i)} + \zeta_{\text{hhldsize}(i)} + \varepsilon_i, \tag{11}$$

where Y_i is time spent on domestic work in minutes per day. The specification is interacted in appliance ownership up to third order, which matters because appliances are unlikely to be additively separable in home production: the time a dishwasher saves is smaller in a household that already owns a clothes washer if the two displace overlapping tasks, and larger if the household's cooking technology makes both more usable. Imposing additivity would attribute to each appliance an average effect that holds at no observed bundle. We stop short of full saturation, however: a fully saturated structure requires $2^{|S|} - 1$ terms, and the highest-order terms are then estimated off very thin and highly correlated cells. Main effects are always retained and only interactions above third order are dropped. Alongside the ownership terms we control for the household's position in the national income distribution and absorb age, country, survey-year, and household-size effects. The regression is weighted by pooled survey weights, rescaled so that each country enters in proportion to its population, and standard errors are clustered by country. Specifications are estimated separately by demographic group, so that the coefficients applied below are those of the relevant population rather than a pooled average.

We then compute, for each unit c, the counterfactual level of domestic work that would

prevail if ownership converged to the rates observed in low-income countries:

$$\widehat{Y}_c^{\mathrm{cf}} = \bar{Y}_c + \sum_{\tau \subseteq S} \widehat{\beta}_\tau \big(x_\tau^{\mathrm{cf}} - \bar{x}_{c,\tau}\big), \tag{12}$$

where τ indexes the terms appearing in (11) — singletons, pairs, and triples — with $\widehat{\beta}_\tau$ the corresponding estimated coefficient. Each $\bar{x}_{c,\tau}$ is the weighted *joint* ownership mean of the appliances in τ, not a product of marginals, so the observed correlation between appliances is carried directly into the calculation; x_τ^{cf} is the corresponding quantity in the lowest income group, taken as a population-weighted average. Summing over subsets rather than over individual appliances is what carries the interaction structure through to the counterfactual. Because the model is linear in the ownership terms, all fixed effects and covariates cancel in the difference, so (12) is an Oaxaca–Blinder-style reweighting: it holds the estimated returns to each bundle fixed and varies only the endowment. Confidence intervals follow from the stored coefficient covariance matrix by the delta method, using a t multiplier with degrees of freedom given by the number of countries in the source regression. For the income-group figures below, ownership shares are averaged within a group first and a single effect computed from the group-level profile, rather than averaging country-level effects: since every country shares the same estimated coefficients, treating country effects as independent draws would understate the group's true uncertainty. Appendix A.4 gives the full construction.

Two caveats follow from this design, and both cut in identifiable directions. First, because (11) includes country effects, the $\widehat{\beta}_\tau$ are within-country returns to appliance ownership, while (12) applies them to between-country differences in ownership rates. If the households that own a given bundle within a poor country differ from the average household in a rich country in ways not captured by income rank, age, and household size, the counterfactual will be biased, and the direction is not obvious a priori. Second, coverage is bought with imputation. A country is otherwise dropped from a set's regression whenever any one appliance in that set is unobserved there, even when the outcome and every other appliance are measured, so we fill whole-column gaps from external national ownership rates and, where those are missing, from a fitted logistic curve in log GDP. This widens the sample considerably but treats imputed ownership as though observed, which makes imputation-heavy estimates optimistically precise. We therefore read the resulting figure as an accounting exercise rather than as a causal estimate of what removing appliances would do.

Figure 25 presents the result. In low-income countries the counterfactual is by construction almost identical to the actual level. In lower-middle-income countries the adjustment is small, at around four minutes per day. It grows to roughly 23 minutes in upper-middle-income

countries and to almost one hour in high-income countries, where counterfactual domestic work is 4.1 hours per day against an actual 3.1. Comparing the actual decline from the lower-middle-income to the high-income group (3.0 hours per day) with the counterfactual decline (2.1 hours), the diffusion of these four appliances accounts for around 30% of the total.

[Figure 25]

Thirty per cent is a large share for a handful of appliances, and it is a lower bound on the technology channel as a whole for two reasons. It omits the infrastructure that makes appliances useful and that itself substitutes for domestic labour — piped water, electrification, gas for cooking — and it omits market substitutes for home production such as prepared food, cleaning services, and laundry. There is also one force pushing the other way, though we expect it to be small: because ownership of an unobserved appliance is imputed independently of the appliances we do observe, and appliances in fact co-move on a common modernisation factor, joint ownership is understated and with it any contribution from positive interaction terms. On balance, then, thirty per cent is likely conservative. It is also, however, only a share: the majority of the decline in domestic work over development remains unaccounted for by appliance access alone. Establishing what fills that gap is the natural next step, and one that our infrastructure is well placed to support given that the underlying diaries retain the full spell structure of domestic activity.

6 Conclusion

This paper constructs a global harmonised time-use database covering more than 210 surveys across 80 countries and uses it to document what parenthood does to the full allocation of time. The existing child-penalty literature measures these costs almost exclusively through paid work. We show that restricting attention to the labour market misses dimensions of the child penalty that are at least as large and more consequential for welfare. We have presented three main sets of findings.

First, the time costs of parenthood are both larger and more heterogeneous than employment-based measures suggest. The arrival of a first child imposes around four additional hours of active time per day on the couple, financed entirely by leisure including sleep, and this cost is persistent: it has not dissipated fifteen years after birth. The aggregate cost intensifies sharply with development — parental childcare roughly triples between low- and high-income countries — and undergoes a qualitative transformation from physical survival tasks in low-

income settings to play, education, and developmental investment in richer ones. Neither of the two mechanisms that might have absorbed this rising cost does so. The broad alloparental care networks that absorbed around 30% of all childcare at the bottom of the income distribution have almost entirely disappeared in advanced economies, with the decline driven overwhelmingly by young alloparents; and formal schooling, which expands enormously over the same range, turns out to complement rather than substitute for parental time, adding around 24 minutes of daily parental childcare per additional hour of school access for children under five. High-income parents therefore face not only more childcare in total but a higher share of it falling exclusively on themselves, even as institutional care expands around them.

Second, domestic work matters as much as childcare in generating the gap between mothers and fathers. We show that unpaid domestic work surges for women at childbirth and contributes to the total gender divergence in active time as much as childcare does. Our task-allocation model shows that the only way to rationalise the observed degree of gender specialisation is a perceived-productivity ratio of mothers to fathers exceeding five — a magnitude that points to norms and gender-stereotyped beliefs rather than biological comparative advantage. Independent evidence supports that reading: across countries, the gender gap in unpaid work is tightly correlated with directly elicited beliefs about whether a pre-school child suffers when the mother works. The same lens also reframes the global regularity, since the couples whose days diverge most at childbirth are those in the regions where they were most alike beforehand: children add least to gender specialisation where it is already greatest. The welfare cost extends further still: average physical effort rises by around 20% for women at birth and remains persistently elevated, while men's physical-effort profiles are essentially flat. Standard child-penalty measures based on employment or earnings are therefore silent on a welfare divergence that operates through the character of work, not just its duration.

Third, the cross-country and within-household evidence together explain why gender specialisation persists and what it would take to undo it. Over the development path, the recovery of female labour force participation at high income is associated with a large decline in maternal domestic work, not a decline in childcare — parental childcare continues to rise throughout. Policies that expand the availability or reduce the cost of market childcare may therefore have limited effects on female labour supply if parental and market childcare are complements rather than substitutes. Within the household, the task complementarity evidence shows why the domestic-work burden is so hard to dislodge: when mothers shift toward paid work, fathers do not absorb more domestic work or childcare. Households do not rebalance symmetrically. This asymmetry holds across income levels and all six coun-

tries in our sample, and it implies that the path to greater gender equality in time use runs through changes in the production technology of domestic work — appliances, infrastructure, market substitutes — rather than through changes in wages or childcare policy alone. We provide a first quantification of that channel: holding household appliance ownership at low-income-country rates would undo around 30% of the observed decline in domestic work between lower-middle-income and high-income countries. That leaves the majority of the decline still to be explained, which we regard as one of the more promising open questions the data raise.

To conclude, by documenting these patterns across settings with fundamentally different institutions and levels of development, our analysis allows for richer investigation of mechanisms than within-country analyses of a single institutional setting. Identifying and quantifying the forces that make the domestic-work burden so hard to dislodge — whether normative, related to perceived productivity differences, or reflecting coordination failures — is the central next step for the economics of gender inequality and household task allocation. We hope the data and descriptive architecture assembled here will provide a foundation for that work.

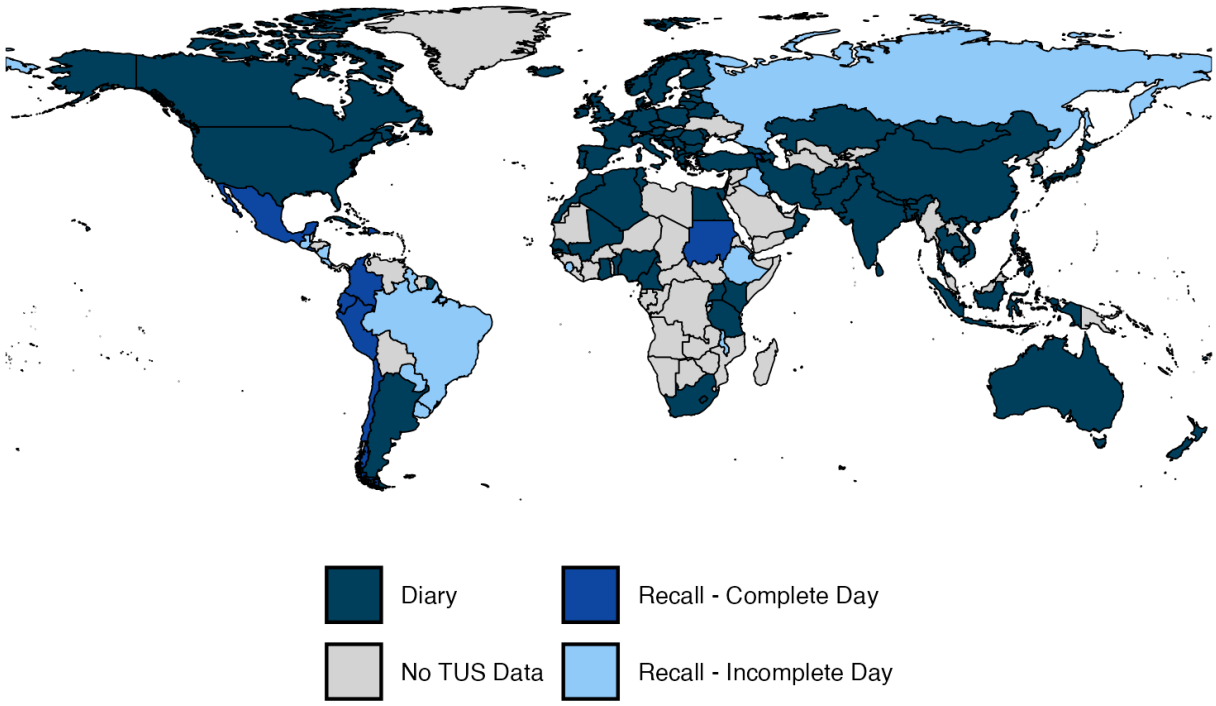


Figure 1: Map of Data Sources

Notes: Countries are classified as diary-based, recall (summing to 24 hours), incomplete recall, or no data. The database covers more than 210 surveys across 80+ countries, representing approximately 90% of the world population.

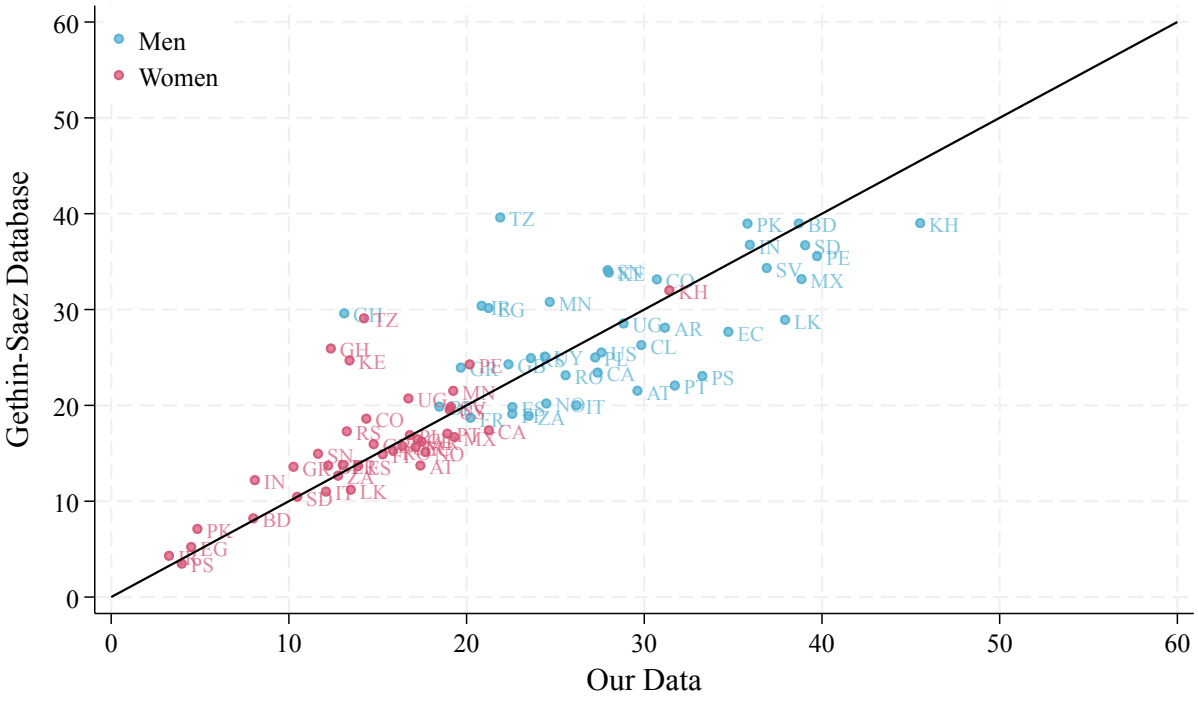
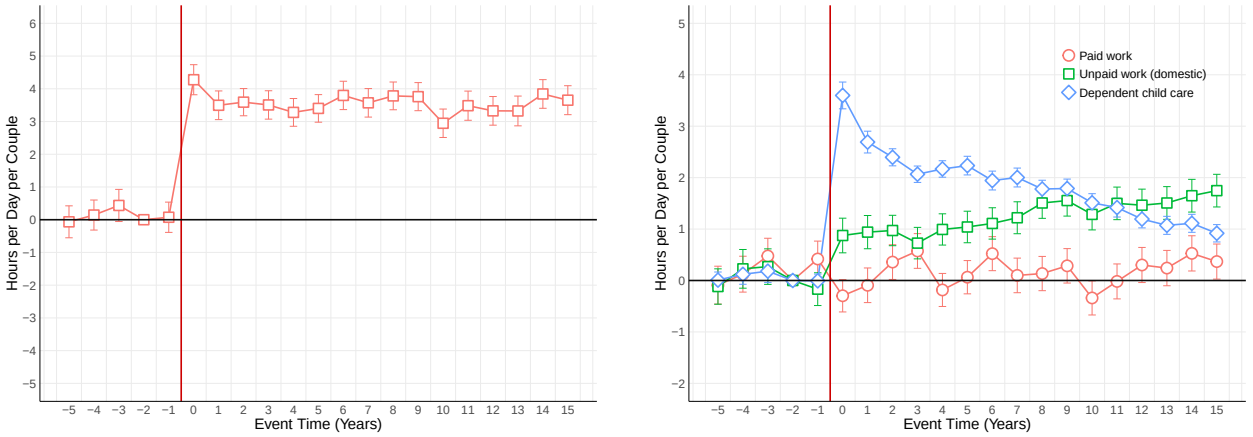


Figure 2: Validation of paid work hours against labour-force survey data

Notes: The horizontal axis shows average weekly hours of paid work in our time-use database; the vertical axis shows the corresponding estimate from Gethin and Saez (2025). Men and women are shown in separate colours. The 45-degree line is shown for reference. Deviations from the 45-degree line are driven by differences in survey timing and seasonality.

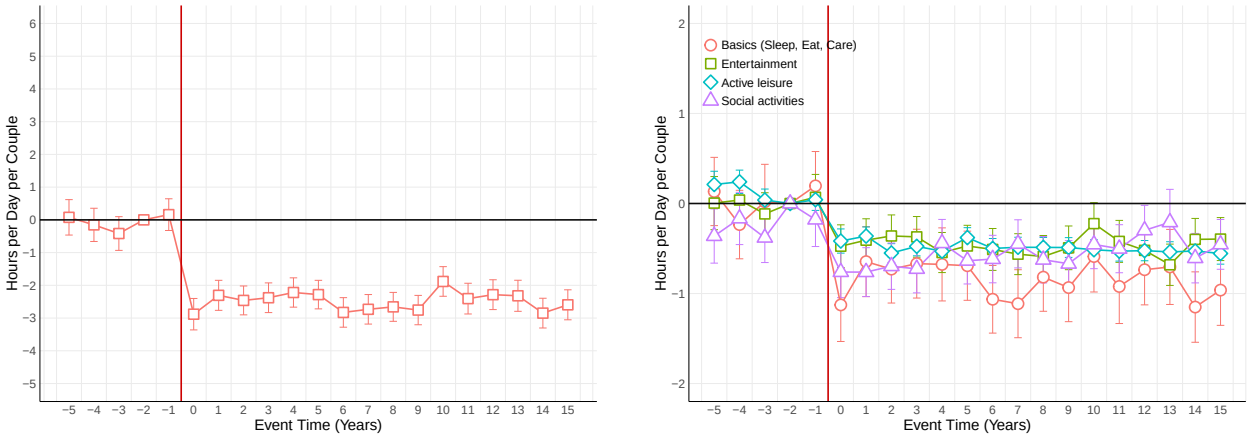


(a) Total active time

(b) Decomposition of active time

Figure 3: Effect of first birth on active time

Notes: Couple-level event studies around first childbirth, pooling all countries. Active Time is defined as the sum of paid work, unpaid domestic work, care work, production for own use and work commute. Panel A shows total active time. Panel B decomposes active time into its main components: dependent childcare, unpaid domestic work, production of goods for own use, and paid work. The horizontal axis shows years relative to the birth of the first child; the vertical axis shows minutes per day relative to $t = -2$. Bands show 95% confidence intervals.

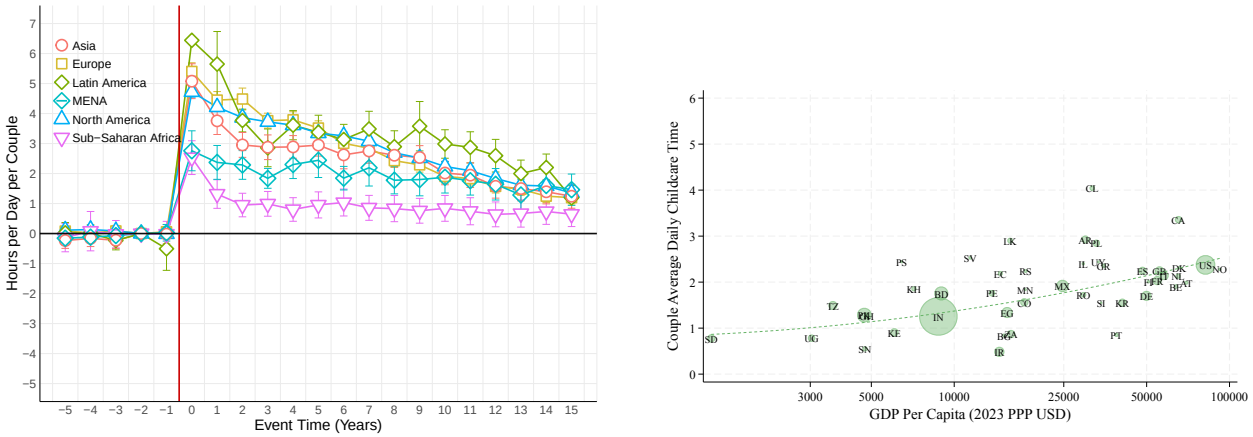


(a) Total Leisure

(b) Decomposition of Leisure

Figure 4: Effect of first birth on leisure

Notes: Couple-level event studies for Leisure around first childbirth, pooling all countries. Panel A shows total Leisure. Panel B disaggregates into the four two-digit components: basics (sleep, eating, personal care), entertainment, active leisure, and social activities. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals.



(a) Event study for childcare by world region

(b) GDP per capita and couple childcare time

Figure 5: Childcare investment over development

Notes: Panel A presents couple-level event studies for childcare around first childbirth for six world regions. Panel B presents a scatterplot of GDP per capita (2023 PPP USD, horizontal axis) against couple daily childcare time (vertical axis); each dot is a country and dot size is proportional to population.

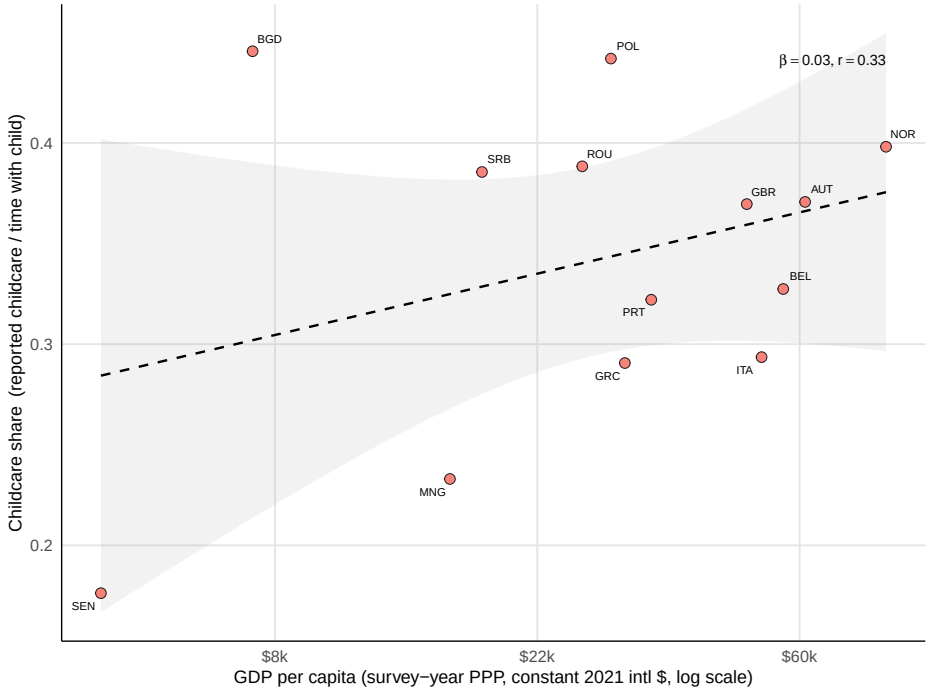
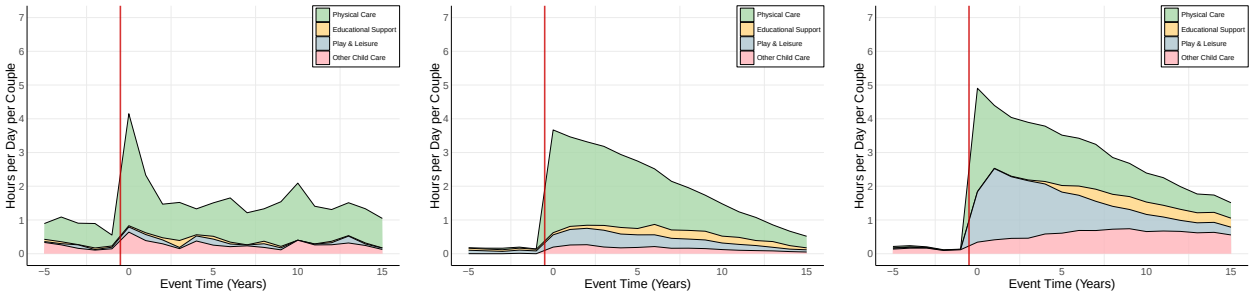


Figure 6: Active childcare relative to time with the child over development

Notes: Each dot is a survey. The vertical axis is the ratio of time reported as childcare to total awake time spent in the presence of the child, for mothers whose oldest child is aged seven or below. The horizontal axis is GDP per capita (survey-year PPP, constant 2021 international dollars, log scale). The dashed line is an OLS fit with a 95% confidence band. The sample is restricted to the surveys that link co-presence to specific household roster members.



(a) Tanzania (b) India (c) United States

Figure 7: Composition of childcare over development

Notes: Stacked event studies for the composition of childcare around first childbirth in Tanzania (Panel A), India (Panel B), and the United States (Panel C). Childcare is disaggregated into physical care, educational support, play and leisure, and other. Each stack shows minutes per day relative to $t = -2$.

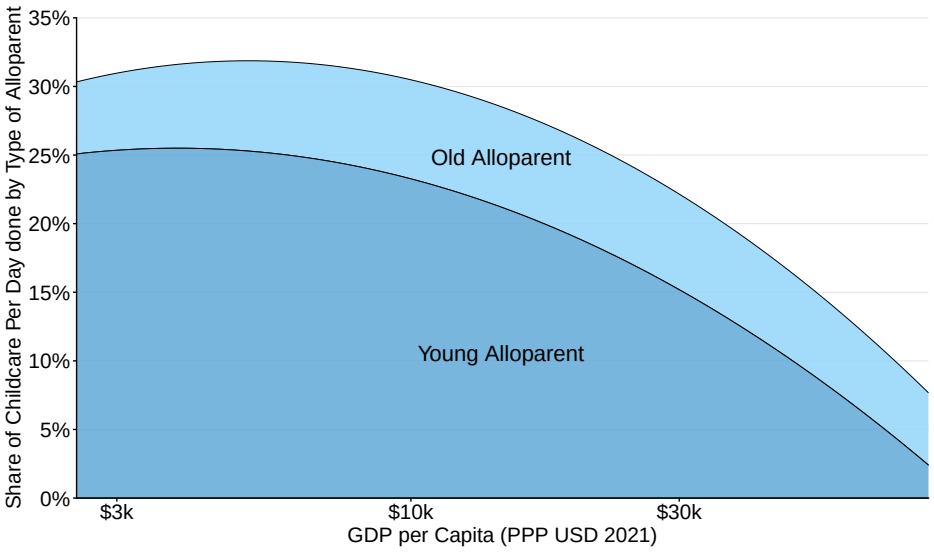


Figure 8: Role of alloparental care over development

Notes: Share of total daily childcare performed by alloparents, plotted against GDP per capita (PPP USD 2021, log scale). Childcare is summed over all household members aged 15 and above. Young alloparents are siblings and other non-parent household members; old alloparents are grandparents, extended family members whose own children are grown, and other older household members. The two areas are cumulative, so the upper envelope is the total alloparental share.

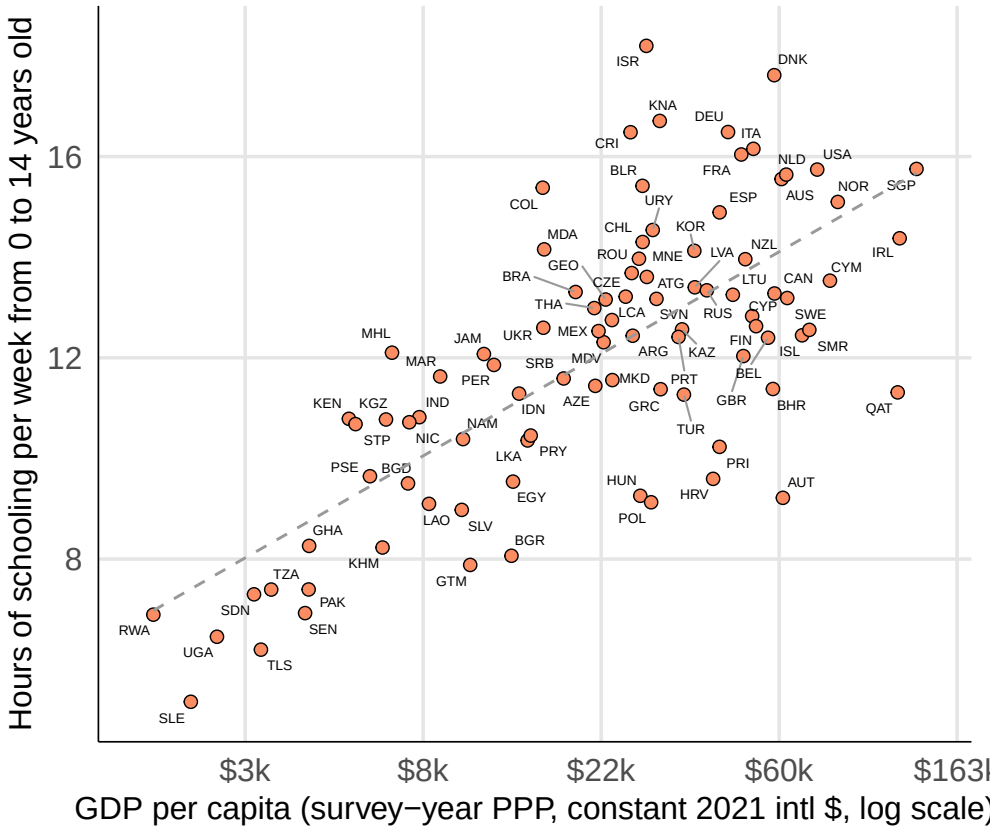
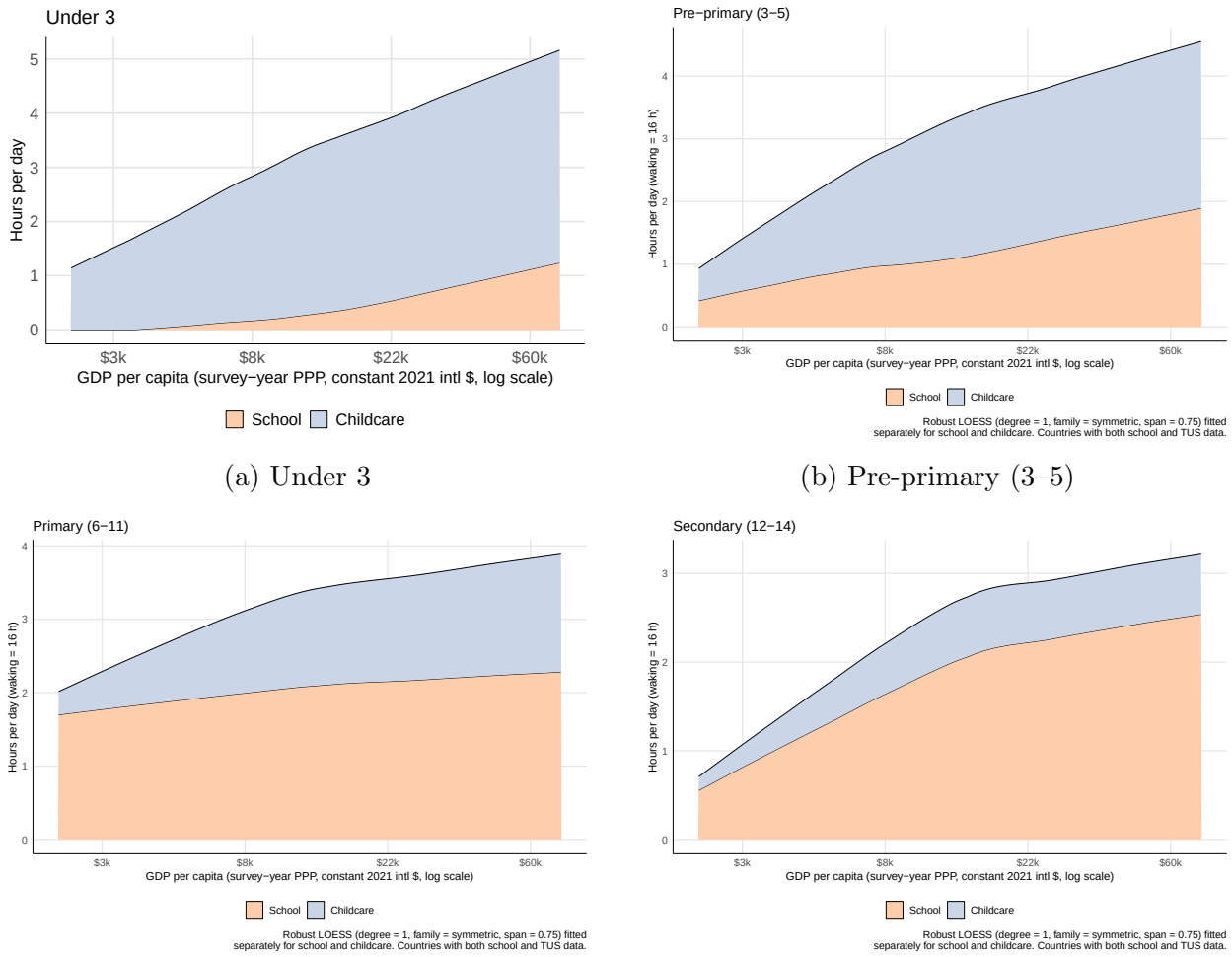


Figure 9: Institutional care over development

Notes: Each dot is a country. The vertical axis is average hours of schooling received per week by children aged 0 to 14, constructed from age-specific enrolment rates, days of instruction per year, and instructional hours per day. The horizontal axis is GDP per capita (survey-year PPP, constant 2021 international dollars, log scale). The dashed line is an OLS fit. 103 countries.



(a) Under 3

(b) Pre-primary (3–5)

(c) Primary (6–11)

(d) Secondary (12–14)

Figure 10: Schooling and parental childcare received, by child age group

Notes: Fitted hours per day of schooling and of received parental childcare, plotted against GDP per capita (survey-year PPP, constant 2021 international dollars, log scale), for children under 3, aged 3–5, 6–11, and 12–14. Areas are cumulative. Fits are robust LOESS (degree 1, symmetric family, span 0.75) estimated separately for schooling and childcare. In Panel A schooling is fitted on all 103 countries and childcare on the 33 countries with time-use data; Panels B to D restrict both fits to the countries with both schooling and time-use data. A waking day is normalised to 16 hours. Received childcare is constructed from parent-reported childcare as described in Appendix A.3.

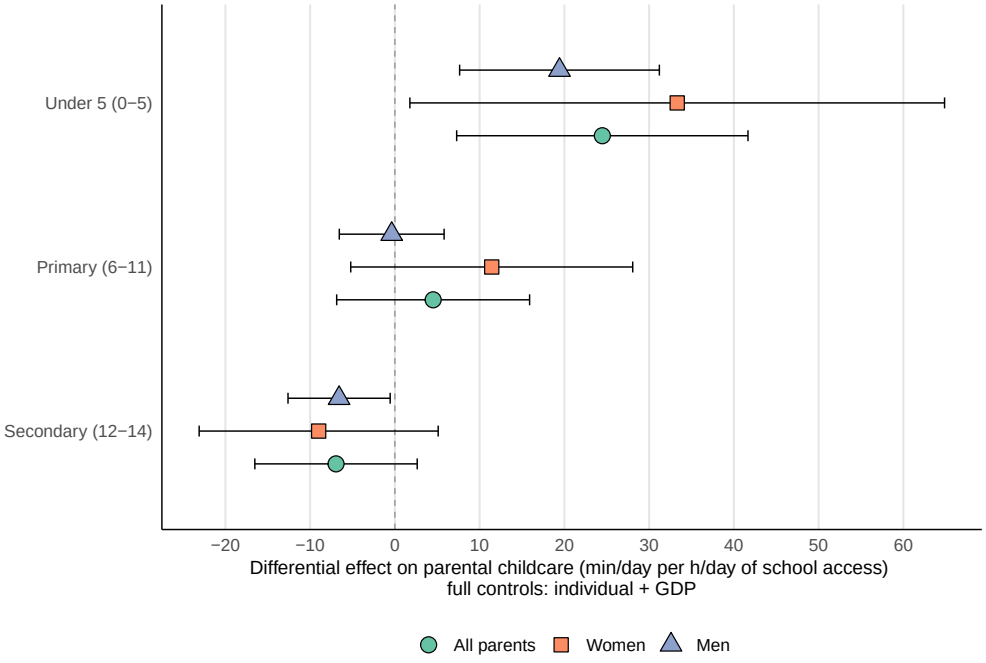
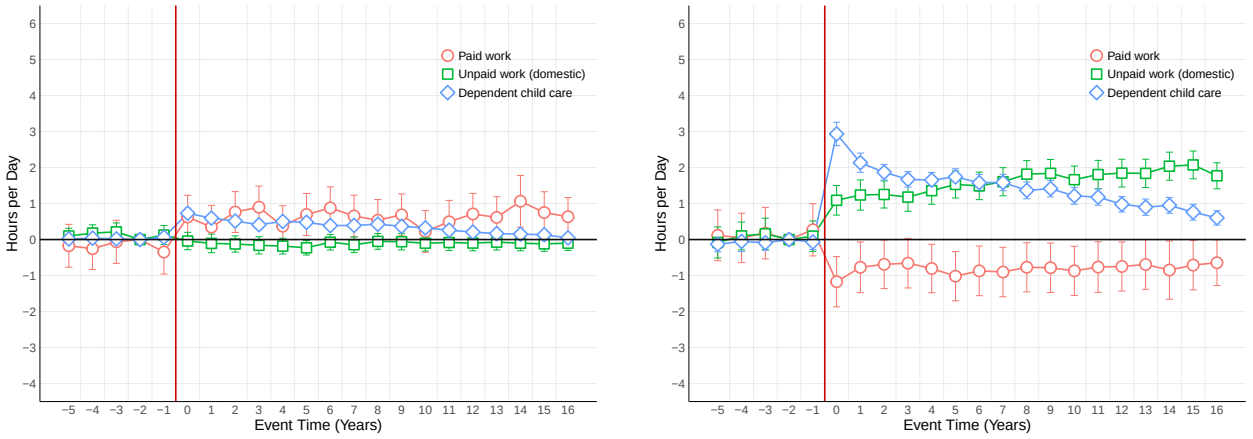


Figure 11: Differential response of parental childcare to school access, by child age group

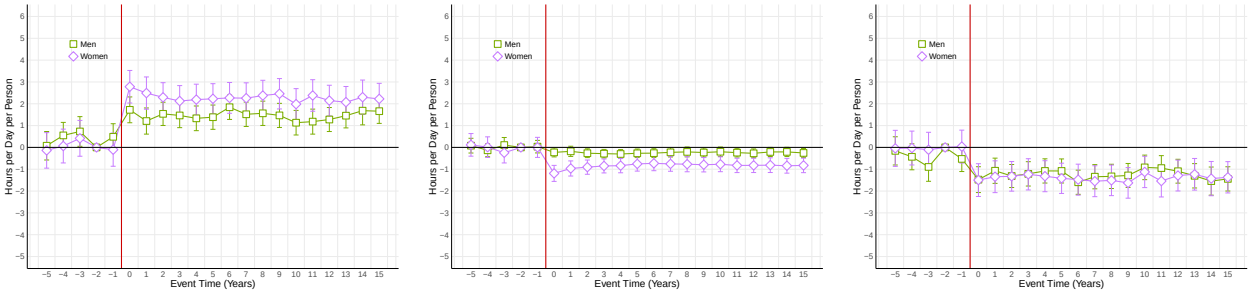
Notes: Estimated interaction coefficients γ_k from equation (2), in minutes per day of parental childcare per additional hour per day of school access, for parents with a child in the indicated age group relative to parents whose children are all in other age groups. Estimates are shown for all parents and separately for mothers and fathers. Controls: age, age squared, ever married, number of children, education fixed effects, and GDP per capita. 26 countries, population-adjusted weights, standard errors clustered by country. Bars show 95% confidence intervals.



(a) Men (b) Women

Figure 12: Effect of first birth on time use by gender

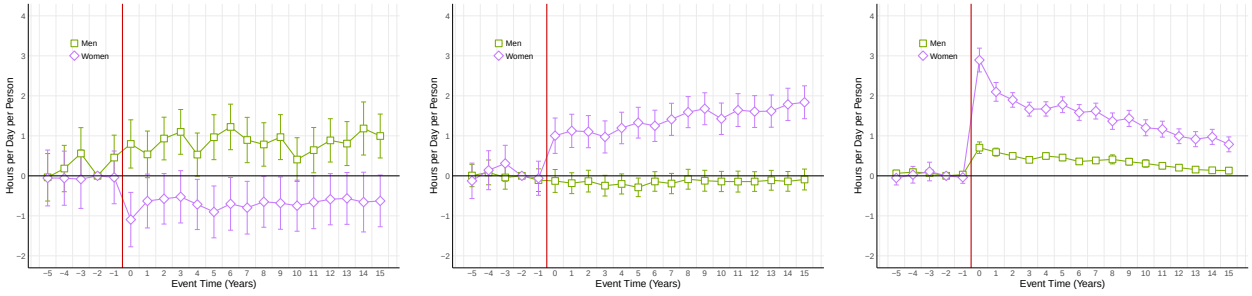
Notes: Event studies around first childbirth for men (Panel A) and women (Panel B) in the same couple, showing paid work, unpaid domestic work, and childcare. The vertical axis shows minutes per day relative to $t = -2$. Estimates pool all countries.



(a) Active time (b) Education (c) Leisure

Figure 13: Effect of first birth by gender: broad categories

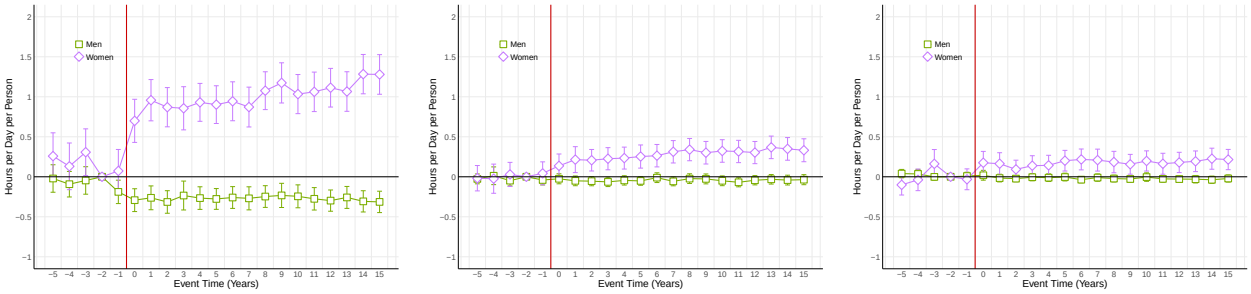
Notes: Event studies around first childbirth for men and women, showing one-digit time-use categories: total active time (Panel A), education (Panel B), and leisure (Panel C). The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates pool all countries.



(a) Paid work (b) Unpaid domestic work (c) Childcare

Figure 14: Effect of first birth by gender: decomposition of active time

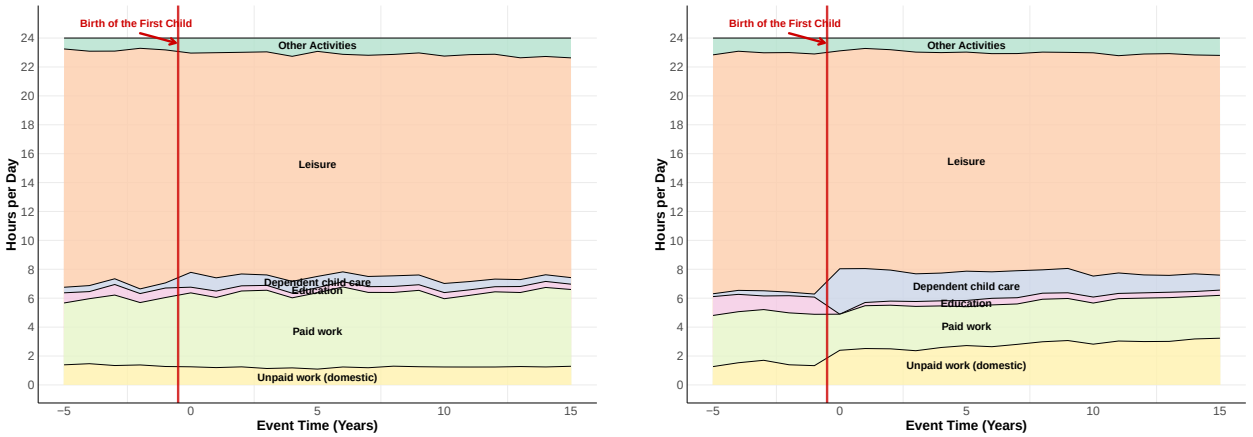
Notes: Event studies around first childbirth for men and women, showing two-digit decomposition of active time: paid work (Panel A), unpaid domestic work (Panel B), and childcare (Panel C). The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates pool all countries.



(a) Food preparation (b) Cleaning (c) Laundry

Figure 15: Effect of first birth by gender: decomposition of domestic work

Notes: Event studies around first childbirth for men and women, showing three-digit components of unpaid domestic work: food preparation (Panel A), cleaning (Panel B), and laundry (Panel C). The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates pool all countries.



(a) Men (b) Women

Figure 16: Full reallocation of the 24-hour day by gender

Notes: Stacked event studies around first childbirth showing the full 24-hour allocation for men (Panel A) and women (Panel B). Each stack shows minutes per day in each time-use category relative to $t = -2$; categories sum to 1,440 minutes per day. Estimates pool all countries.

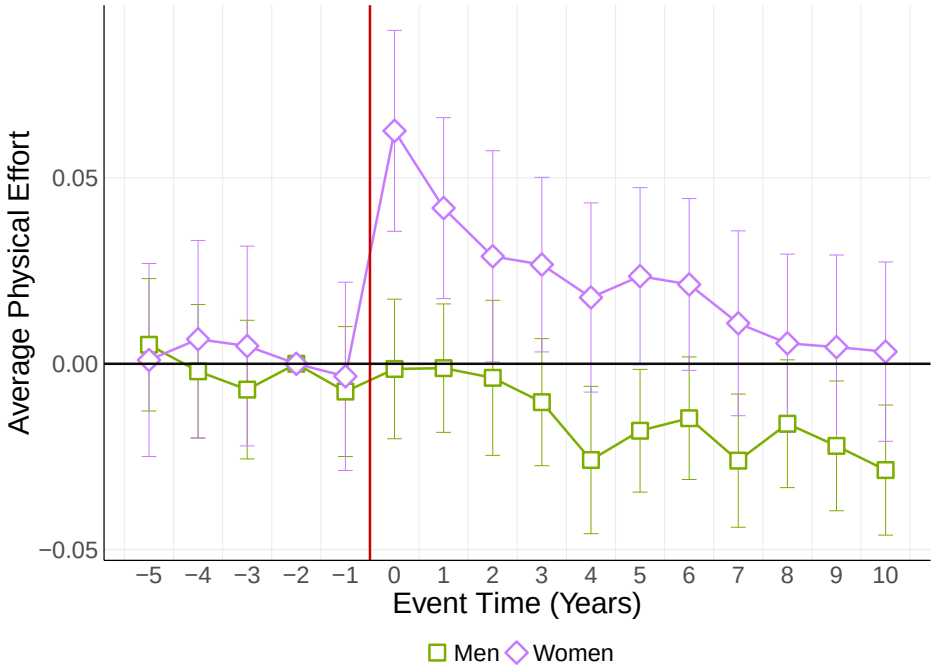


Figure 17: Physical effort around first birth, by gender

Notes: Event study of average daily physical effort around first childbirth, separately for men and women. Physical effort is a continuous score assigned to each activity and aggregated to the individual-day level by taking the mean over the activity portfolio, building on Cornwell et al. (2019). The vertical axis shows the change in the average score relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates pool all countries.

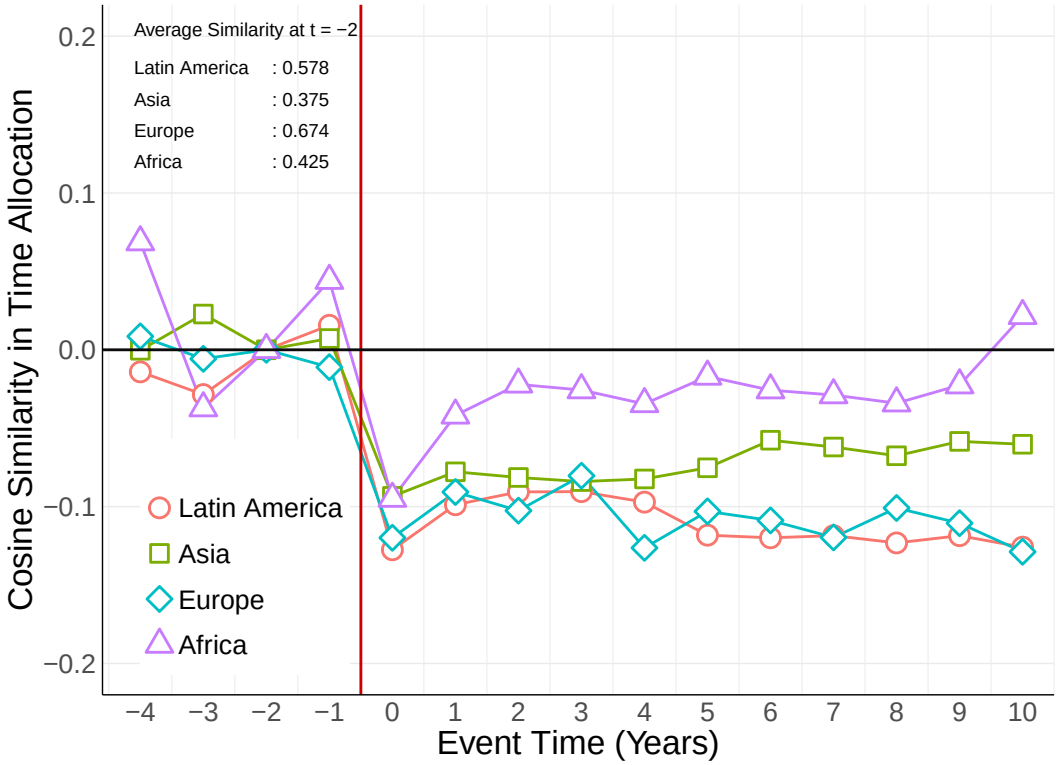


Figure 18: Within-household cosine similarity around first birth, by region

Notes: Event study of within-household cosine similarity of time-use vectors around first childbirth, presented separately for Latin America, Asia, Europe, and Africa. For each couple, cosine similarity is computed between the mother's and father's time-use vectors projected into the full n-dimensional activity space, excluding sleep. A value of one implies identical activity portfolios; a value of zero implies completely orthogonal portfolios. The vertical axis shows the change in similarity relative to $t = -2$; the inset reports the level of similarity at $t = -2$ in each region.

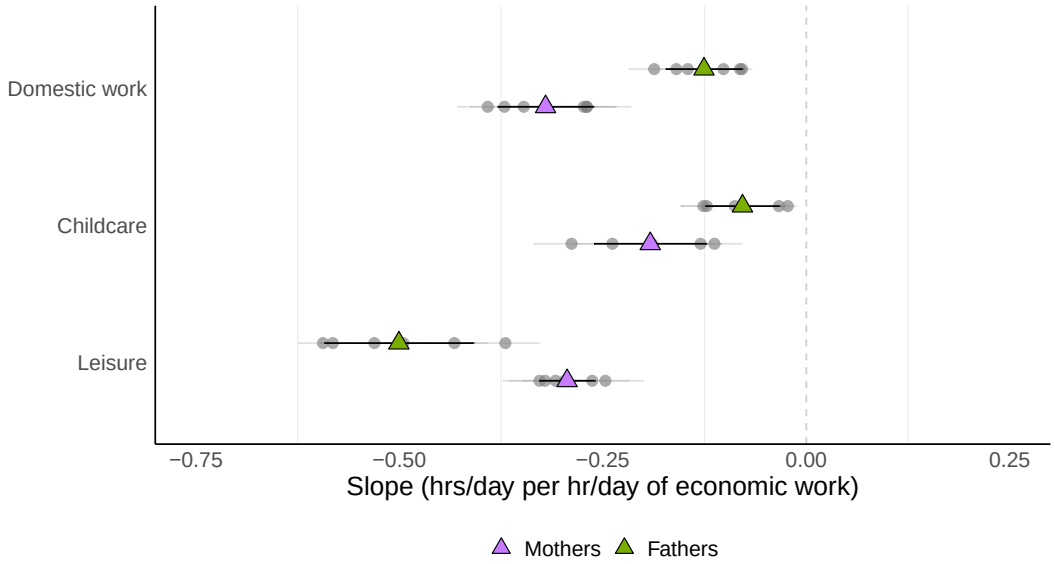


Figure 19: Within-person task complementarities, by country and gender

Notes: Within-person slope coefficients (equation 4) from a binscatter of residualised time use on residualised economic work hours, for six countries spanning the income distribution and for the cross-country average. Each coefficient shows the substitution between a parent's own economic work and their own domestic work, childcare, and leisure, separately for mothers and fathers. Estimates condition on age, education, marital status, household size, number of children, and day of week. Sample: couple households with youngest child ≤ 10 years old.

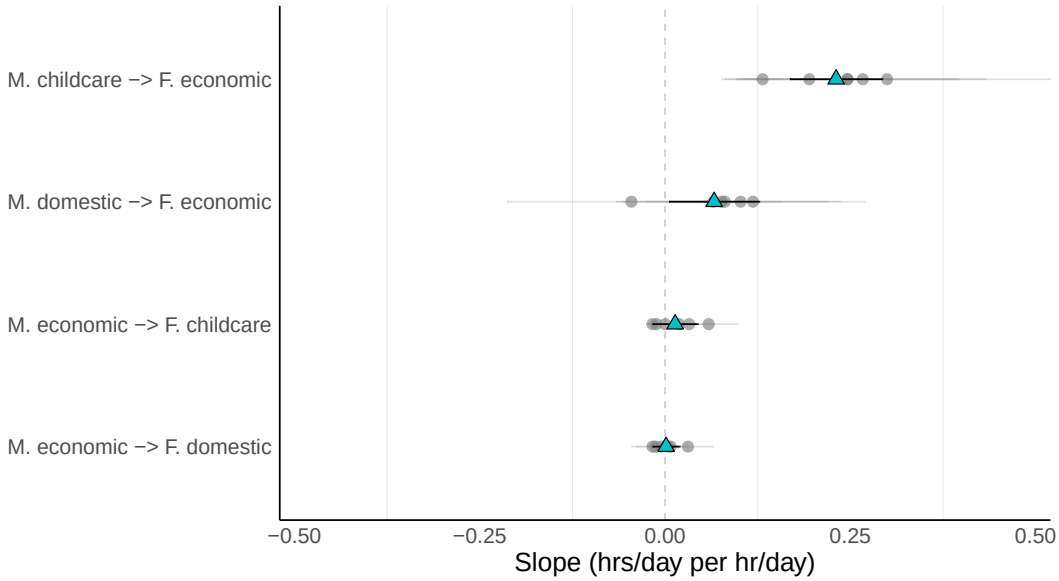
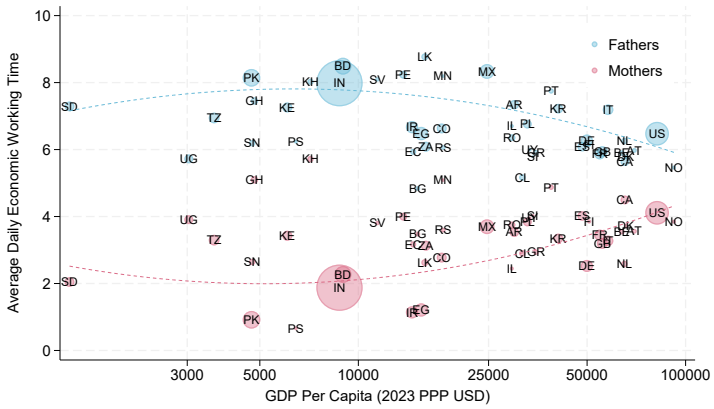
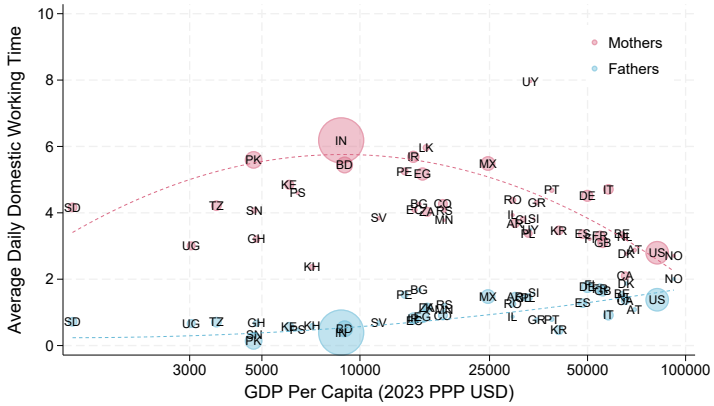


Figure 20: Cross-spouse task complementarities, by country

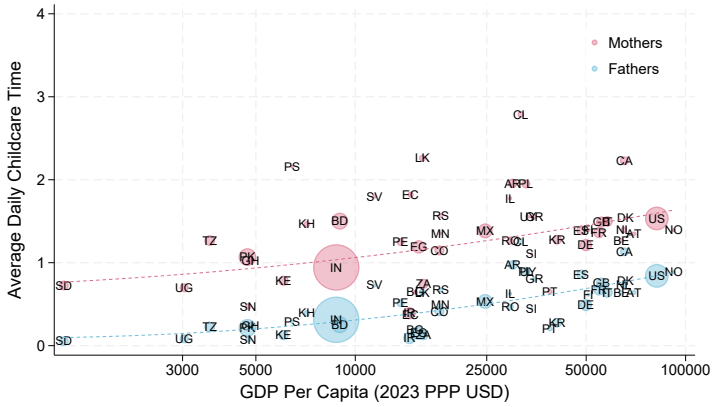
Notes: Cross-spouse slope coefficients (equation 5) from a binscatter of residualised time use on the partner's residualised economic work hours, for six countries spanning the income distribution and for the cross-country average. The four pairs shown are (i) mother's childcare → father's economic work, (ii) mother's domestic work → father's economic work, (iii) mother's economic work → father's childcare, and (iv) mother's economic work → father's domestic work. Estimates condition on age, education, marital status, household size, number of children, and day of week. Sample: couple households with youngest child ≤ 10 years old.



(a) Paid work



(b) Domestic work



(c) Childcare

Figure 21: Parenthood and development: time allocations of mothers and fathers

Notes: Scatterplots of average daily time allocations for mothers and fathers against GDP per capita (2023 PPP USD), for paid work (Panel A), domestic work (Panel B), and childcare (Panel C). Each dot is a country; dot size is proportional to population. Mothers and fathers are shown in different colours.

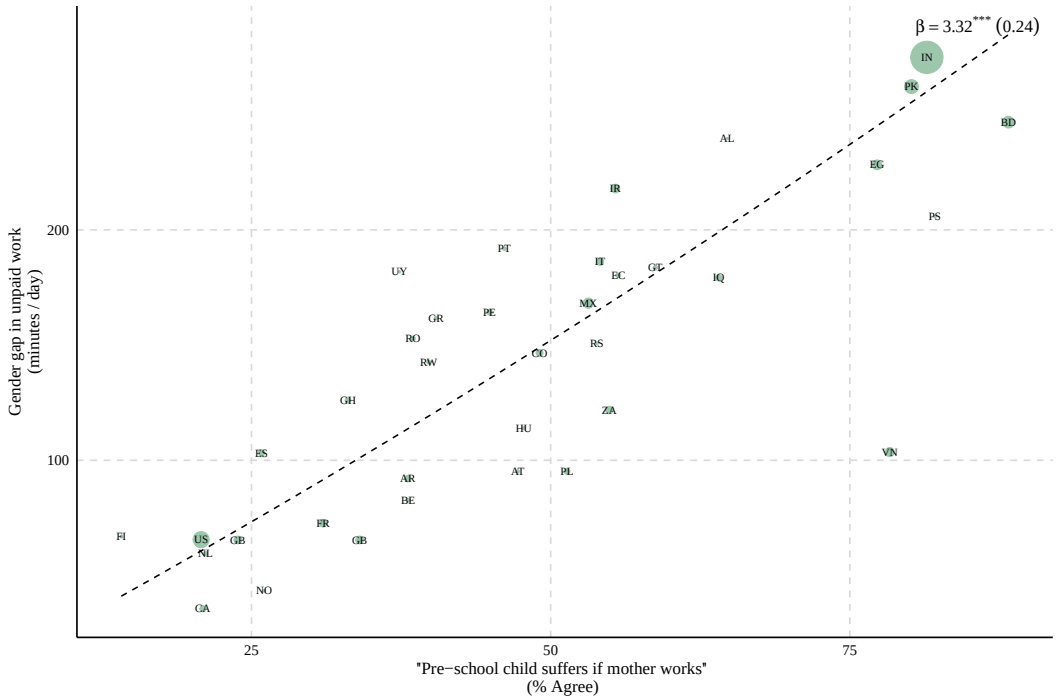


Figure 22: Elicited gender-role norms and the gender gap in unpaid work

Notes: Each dot is a country; dot size is proportional to population. The vertical axis is the gender gap in unpaid work, defined as mothers' minus fathers' average daily minutes. The horizontal axis is the share of survey respondents agreeing that a pre-school child is likely to suffer if the mother works. The dashed line is a population-weighted OLS fit; $\beta = 3.32$ (0.24).



Figure 23: Domestic work declines with development

Notes: Average daily hours of domestic work by World Bank income group, pooling all surveys in our database.

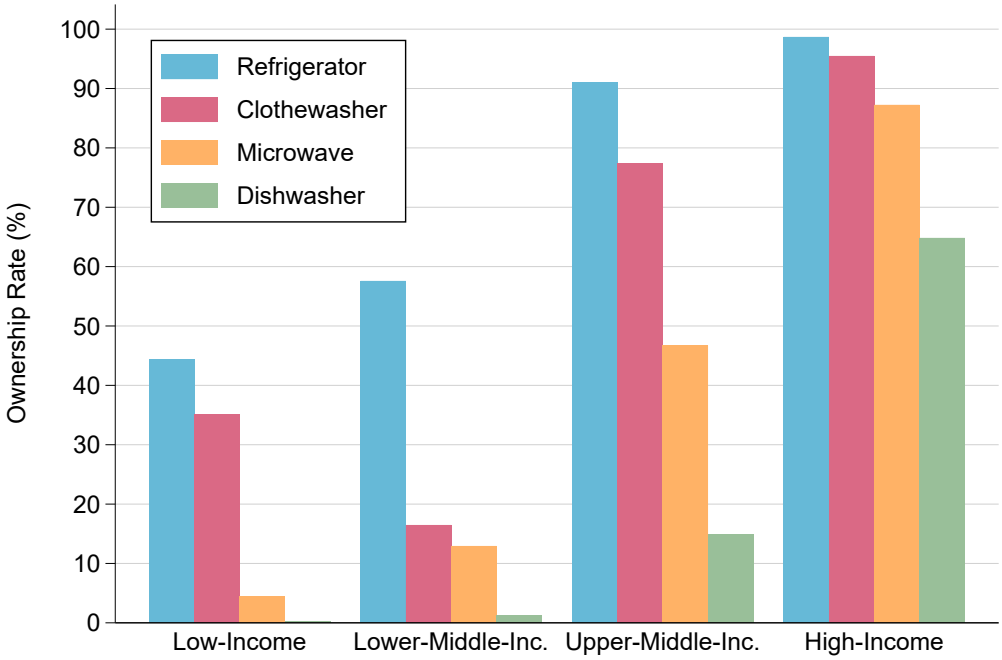


Figure 24: Household appliance ownership rises with development

Notes: Share of households owning each appliance, by World Bank income group, in the surveys that record both appliance ownership and time use.



Figure 25: Domestic work: actual and counterfactual at low-income appliance ownership

Notes: Actual average daily hours of domestic work by income group, alongside the counterfactual from equation (12) holding appliance ownership at the population-weighted rates of the lowest income group. Estimated on the surveys that record both appliance ownership and time use, widened by the ownership imputation described in Appendix A.4. Confidence intervals, where shown, are delta-method intervals computed from the coefficient covariance matrix.

References

Aguiar, Mark and Erik Hurst, "Measuring Trends in Leisure: The Allocation of Time over Five Decades," *Quarterly Journal of Economics*, 2007, *122* (3), 969–1006.

Aguilar-Gomez, Sandra, Eva O. Arceo-Gomez, and Elia De la Cruz Toledo, "Inside the black box of child penalties: Unpaid work and household structure," *Journal of Development Economics*, 2026, *178*, 103554.

Becker, Gary S., "An Economic Analysis of Fertility," in "Demographic and Economic Change in Developed Countries," National Bureau of Economic Research / Columbia University Press, 1960, pp. 209–240.

Cornwell, Benjamin, Jonathan Gershuny, and Oriel Sullivan, "The Social Structure of Time: Emerging Trends and New Directions," *Annual Review of Sociology*, 2019, *45*, 301–320.

Cowan, Ruth Schwartz, *More Work for Mother: The Ironies of Household Technology from the Open Hearth to the Microwave*, New York: Basic Books, 1983.

Cuevas-Ruiz, Pilar, José Ignacio Gimenez-Nadal, Sveva Manfredi, and Almudena Sevilla, "A new model of parental time investments: a paradigm shift for addressing gender inequality in the labour market," *Oxford Review of Economic Policy*, 2026, *41* (3-4), 820–842.

Cunha, Flavio and James J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

de Grignon Perez, Juana Lamote, Camille Landais, Tanay Kasyap, Jiaqi Tang, Jean Fontalirand, Gabriel Leite-Mariante, Eloïse Leroux, Sana Rashidi, and Jonathan Gershuny, "Multinational Time Use Study Plus (MTUS+), Version [N]," Technical Report, Centre for Time Use Research, University College London, London 2026.

Dinkelman, Taryn and L. Rachel Ngai, "Time Use and Gender in Africa in Times of Structural Transformation," *Journal of Economic Perspectives*, 2022, *36* (1), 57–80.

Doepke, Matthias and Fabrizio Zilibotti, *Love, Money, and Parenting: How Economics Explains the Way We Raise Our Kids*, Princeton, NJ: Princeton University Press, 2019.

_ **and Michèle Tertilt**, "Women's Empowerment, the Gender Gap in Desired Fertility, and Fertility Outcomes in Developing Countries," *AEA Papers and Proceedings*, 2018, *108*, 358–362.

Gethin, Amory and Emmanuel Saez, "Global Working Hours," 2025, (34217), 1–40.

Gottlieb, Charles, Cheryl Doss, Douglas Gollin, and Markus Poschke, "The Gender Division of Work across Countries," IZA Discussion Paper 16896, IZA Institute of Labor Economics 2024.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu, "Engines of Liberation," *The Review of Economic Studies*, 2005, *72* (1), 109–133.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney, "Parental Education and Parental Time with Children," *Journal of Economic Perspectives*, 2008, *22* (3), 23–46.

Hrdy, Sarah Blaffer, *Father Time: A Natural History of Men and Babies*, Princeton, NJ: Princeton University Press, 2024.

International Labour Organization, "Global Wage Report 2024–25: Is Wage Growth Bouncing Back?," Technical Report, International Labour Organization, Geneva 2024.

Kleven, Henrik, "The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach," Working Paper 30176, National Bureau of Economic Research 2022.

_ **, Camille Landais, and Gabriel Leite-Mariante**, "The Child Penalty Atlas," *Review of Economic Studies*, 2025, *92* (5), 3174–3207.

_ **,** _ **, and Jakob Egholt Søgaard**, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019, *11* (4), 181–209.

Lamb, Michael E., ed., *The Role of the Father in Child Development*, 5th ed., Hoboken, NJ: John Wiley & Sons, 2010.

Ngai, L. Rachel and Barbara Petrongolo, "Gender Gaps and the Rise of the Service Economy," *American Economic Journal: Macroeconomics*, 2017, *9* (4), 1–44.

_ **, Claudia Olivetti, and Barbara Petrongolo**, "Structural Transformation over 150 Years of Women's and Men's Work," Working Paper, London School of Economics 2024.

Rilling, James K., *Father Nature: The Science of Paternal Potential*, Cambridge, MA: MIT Press, 2024.

Yeung, W. Jean, John F. Sandberg, Pamela E. Davis-Kean, and Sandra L. Hofferth, "Children's Time with Fathers in Intact Families," *Journal of Marriage and Family*, 2001, *63* (1), 136–154.

A Harmonisation and Measurement

A.1 LLM-assisted activity classification

Mapping the activity codes of each national statistical office onto the three-level hierarchy of Table A1 is the single largest fixed cost of building the database. Classification systems differ across countries and across waves within countries, vary in granularity from a few dozen to several hundred codes, and are documented in a dozen languages. Manual harmonisation is feasible for a handful of countries but does not scale to more than 210 surveys, and it is not costless in accuracy either: human coders disagree with one another on ambiguous labels, and their disagreements are neither documented nor reproducible.

We therefore use a large language model to assign the hierarchical labels, with three features designed to control the failure modes of that approach. First, the pipeline is run as an ensemble: each activity label is classified several times and the modal label is retained, which prevents any single anomalous completion from propagating into the data. Second, the model is constrained to the admissible label set by retrieval over the classification manual, so that it selects among existing categories rather than generating new ones. Third, the output is reviewed and corrected by hand for every country, with the ensemble's disagreement rate used to prioritise which labels receive attention.

Classification runtime falls by orders of magnitude relative to manual coding. After fine-tuning, agreement between the model and human annotators is comparable to agreement between human annotators, and residual disagreement is concentrated in activities that are intrinsically ambiguous rather than distributed across the hierarchy — for instance, activities that could plausibly be classified as either childcare or domestic work, such as preparing food for a child. This is consistent with a growing literature documenting that LLM classification of short text matches expert judgement in comparable settings (**??**).

A.2 From activities to attributes

Activity labels tell us what someone was doing; they do not directly tell us what their day was like. Two parents with identical activity portfolios may differ in how physically demanding, how cognitively loaded, or how spatially confined their days are. Following Cornwell et al. (2019), we extract attributes of activities as a second layer on top of the classification, and we use the physical-effort attribute in Section 5.1.2. The same machinery supports other attributes — cognitive load, location, whether the activity is done alone — which we do not exploit here.

The conceptual framework has three steps. For each attribute a we define a continuous latent variable ν_a and assign a value to every activity in the hierarchy. We then set absolute thresholds ν_a^* that map ν_a into discrete categories where a categorical measure is wanted. Finally we choose an aggregation statistic — mean, median, or maximum — to summarise sub-activity values when constructing scores for higher-order tasks. Three principles guide the design: transparency, in that definitions and cut-points are stated and reproducible; scalability, in that the whole procedure is automated; and flexibility, in that new attributes can be added without touching the underlying microdata.

The scoring itself proceeds in three stages. In the first, all activity labels are sent to the model in a single batch, which forces scores to be assigned relative to one another rather than in isolation and yields a continuous score $\hat{\nu}_a$ per activity. In the second, the batch scoring is repeated and the mean $\bar{\nu}_a$ is retained, which reduces run-to-run variance in the same way as the modal-label ensemble above. In the third, thresholds are set from the data rather than by hand: we run k-means clustering on $\{\bar{\nu}_a\}$ with k equal to the number of desired categories, repeat the clustering, and select the thresholds that minimise the within-cluster sum of squares.

On validation, the procedure achieves an average correlation above 0.7 with human annotators while exhibiting greater internal consistency than the human annotators do — that is, it agrees with itself across runs more closely than independent coders agree with each other, while agreeing with them about as well as they agree among themselves. This mirrors findings elsewhere on LLM-assisted measurement of latent attributes (**???**).

A.3 From parent-reported to child-received childcare

Time-use diaries record childcare from the perspective of the parent, not the child. Fact 4 requires the opposite perspective, because schooling is measured as hours received by a child of a given age. Converting parent-reported care into child-received care requires two corrections.

The first is presence. A child receives only the part of a parent's reported childcare during which that particular child is present. In a household with three children, a parent reporting two hours of childcare is not delivering two hours to each. The second is double-counting. Simply adding the mother's and the father's reported care overstates what the child receives whenever the two overlap, so the father's contribution must be counted only when the mother is not simultaneously present. Writing s^{mom} for the share of the mother's childcare time during which the child in question is present, and s^{dad} for the share of the father's childcare

time during which the child is present *and* the mother is absent, received care is

$$\text{received} = CC^{\text{mom}} \times s^{\text{mom}} + CC^{\text{dad}} \times s^{\text{dad}}. \tag{A1}$$

The shares s^{mom} and s^{dad} require co-presence to be linked to specific members of the household roster, which only the United States survey records. We estimate them there as functions of family structure — the child's age, the number of siblings, and household size — and apply the estimated model to every child in every country, mapping parent-level data into child-level data.

What if the presence model does not travel? Family structure is already an argument of the model, so cross-country differences in family size and composition are accounted for by construction. The threat is that presence differs across countries *conditional* on family structure. The direction matters. If presence is lower in poorer countries than the US-estimated model implies, we are overstating received care at the bottom of the income distribution and therefore *understating* the development gradient: the truth would be steeper than what we report. The adverse case is the reverse — presence higher in poorer countries, as one would expect if care there is more often delivered to several children at once — which would flatten the gradient.

Stress test. We take the adverse case to its limit. We impose that presence is 100% in the poorest country in the sample, falling smoothly to the US-estimated level at the top of the income distribution, which is the most hostile assumption the data admit. Figure A1 reports the resulting gradients of received childcare on log GDP per capita, alongside the unadjusted series and the baseline US model. The gradient attenuates but survives for young children: for children under three it falls from roughly eight to roughly five hours per week per log point, and for the combined under-five group from roughly seven to roughly four, remaining bounded away from zero in both cases. For secondary-age children the gradient is small under every specification and becomes indistinguishable from zero under the stress test. The complementarity result of Fact 4, which is concentrated in the under-five group, is therefore not an artefact of the presence adjustment.

A.4 Appliance ownership and domestic work

This appendix documents the pipeline behind Section 5.3.5. It has two stages: estimation of the interacted ownership regression on a country sample widened by imputation, and construction of counterfactuals from the estimated coefficients with delta-method inference.

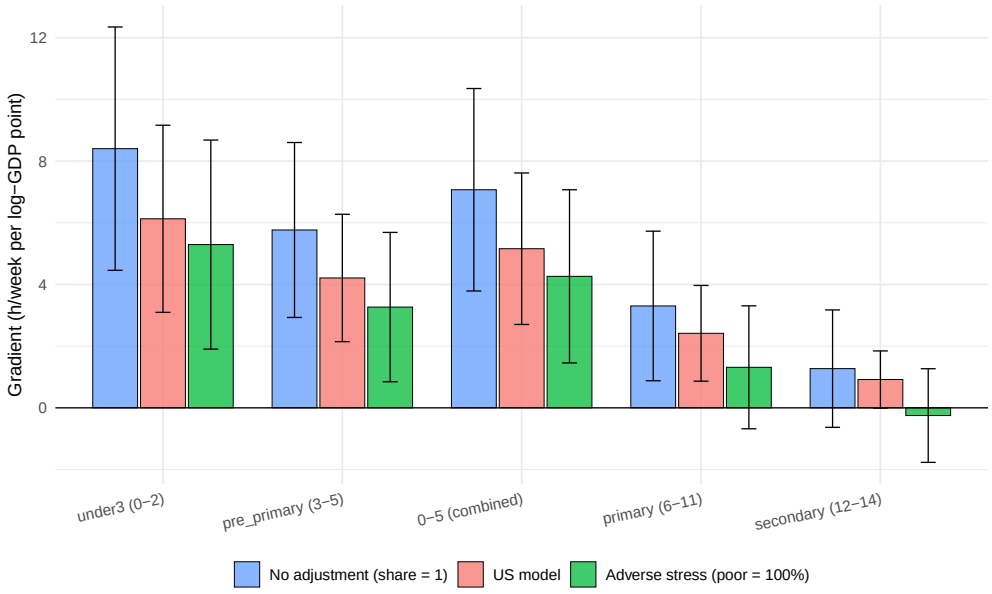


Figure A1: Development gradient in received childcare under alternative presence assumptions

Notes: Bars show the OLS slope of received childcare (hours per week) on log GDP per capita, by child age group, under three assumptions about presence: no adjustment, which assigns all reported parental childcare to every child (share set to one); the baseline presence model estimated on the United States survey; and the adverse stress case, which imposes a presence share of 100% in the poorest country falling to the US-estimated level at the top of the income distribution. Whiskers show 95% confidence intervals.

Objects. The estimation sample is the pooled harmonised microdata, one row per respondent i in country c, carrying a survey weight w_i, binary appliance-ownership indicators $D_{a,ic} \in \{0, 1\}$, and time-use outcomes. Weights are rescaled so that each country enters in proportion to its population. An appliance *set* S is a fixed group of appliances analysed jointly, and a *term* $\tau \subseteq S$ is a single appliance or an interaction of several. We write $\pi_{a,c} \in [0, 1]$ for the national ownership rate of appliance a in country c, a country-level scalar taken from external data or from a fitted curve. Specifications are indexed by demographic group — gender, parenthood status, and whether household income is controlled for — and estimated separately for each.

Estimation. For each combination of outcome, appliance set, and group we fit equation (11) by weighted least squares with fixed effects for age, country, survey year, and household size, and standard errors clustered by country. The right-hand side contains every appliance as a main effect plus interactions up to a configurable order cap, set to three throughout. The cap is a deliberate departure from full saturation: a saturated structure fits one free parameter per distinct ownership pattern, which is $2^{|S|} - 1$ terms and reaches

255 for our largest set, and once the sample is also split by demographic group the highest-order terms are identified off very thin and highly correlated cells. Main effects are always retained; only interactions above the cap are dropped. For each fitted regression we store the coefficients and the full coefficient covariance matrix $\widehat{V} = \widehat{\mathrm{Var}}(\widehat{\beta})$. Storing $\widehat{V}$ is what allows the counterfactual stage to compute standard errors for arbitrary linear combinations of coefficients without re-estimating anything.

Widening the sample. A country is excluded from a set's regression if *any* appliance in that set is entirely unobserved there, even when the outcome and all other appliances are measured. Because many countries with good time-use data are missing a single appliance variable, this exclusion is both costly and arbitrary, so we fill whole-column gaps from external information. Partial gaps within a country are left to the estimator's usual listwise deletion. The national rate $\pi_{a,c}$ is the observed external rate where available, and otherwise the prediction of a per-appliance logistic curve in log income,

$$\mathrm{logit}(\pi_{a,c}) = \gamma_{0,a} + \gamma_{1,a} \log_{10} \mathrm{GDP}_c, \tag{A2}$$

fitted once per appliance on the countries that do have data, with each country given equal weight, and evaluated at each country's own survey-year GDP. Two fill rules are available. The *threshold* rule sets $D_{a,ic} = 0$ where $\pi_{a,c} < 0.05$ and $D_{a,ic} = 1$ where $\pi_{a,c} > 0.95$, leaving intermediate cases missing so that the appliance stays strictly binary; it is the conservative option, and countries with intermediate rates still cannot enter. The *continuous* rule sets $D_{a,ic} = \pi_{a,c}$ for every respondent in the country.

The continuous rule is not circular, for a reason worth stating explicitly. A continuous fill makes the appliance a country constant, and a country-constant regressor is collinear with that country's fixed effect. The imputed country therefore contributes essentially nothing to the imputed appliance's *own* main effect. What it contributes is its rows, which keep the country in the sample so that the other, genuinely observed appliances in the set are estimated on more countries. Interactions between an observed appliance and an imputed one are not country-constant, since they vary with the observed appliance, and so do draw information from the imputed country. The fill buys statistical power for the appliances a country does have rather than manufacturing an effect for the one it lacks.

Counterfactuals. For a term τ, the observed ownership share in country c is the weighted joint mean of the appliances it contains,

$$x_{c,\tau}^{\text{obs}} = \frac{\sum_i w_i \prod_{a \in \tau} D_{a,ic}}{\sum_i w_i}, \tag{A3}$$

which preserves the empirical correlation among appliances rather than assuming independence. Collecting these into a vector x over all fitted terms, the effect of moving country c from its observed profile to a target profile x^{cf} is $\Delta_c = (x^{\text{cf}} - x_c^{\text{obs}})'\widehat{\beta}$, with standard error $\text{SE}(\Delta_c) = \sqrt{d'\widehat{V}d}$ for $d = x^{\text{cf}} - x_c^{\text{obs}}$, and the predicted level is $\bar{Y}_c + \Delta_c$ as in equation (12). Confidence intervals use a t multiplier with degrees of freedom given by the number of clusters in the source regression. Several target profiles are implementable — universal non-ownership or ownership, a single donor country's real joint pattern, a median-GDP basket, ownership read off the fitted curves at a chosen income level — and the scenario used in Section 5.3.5 is the population-weighted average joint ownership of the lowest income bin.

For income-group results, countries are assigned to bins and shares are averaged *within* a bin before a single effect is computed from the bin-level profile. This ordering is deliberate. Every country's effect is computed from the same $\widehat{\beta}$ and $\widehat{V}$, so averaging country-level effects as though they were independent draws would understate the bin's uncertainty; applying the delta method to the bin-level difference d directly is both correct and simpler.

Coverage is widened once more at this stage, by imputing ownership *shares* rather than microdata where an appliance is unobserved. Splitting a term's appliances into the part observed in country c, denoted O_c, and the part that is not, the mixed share is the real empirical joint mean of the observed part times the product of imputed marginals for the missing part:

$$\tilde{x}_{c,\tau} = \overline{\Big(\prod_{a \in \tau \cap O_c} D_{a,ic} \Big)}_{c} \cdot \prod_{a \in \tau \setminus O_c} \pi_{a,c}. \tag{A4}$$

Assumptions and limitations. The pipeline trades rigour for coverage in specific, identifiable places, and we list them in rough order of how much they matter. On estimation: the coefficients are associations net of the fixed-effect structure and are not claimed to be causal effects of ownership; under the continuous fill the imputed appliance's own main effect is essentially unidentified within imputed countries, so its value should be read as what the observed countries say; and because imputed values are treated as observed, the uncertainty of the external rate or the curve prediction is not propagated into the regression standard errors, which makes imputation-heavy estimates optimistically precise. On the ownership

curve: equation (A2) uses income alone, so region, relative prices, and urbanisation are absorbed into residual scatter, and predictions outside the fitting sample's income range are extrapolations; applying a fixed-year external snapshot at each country's own survey year assumes the income–ownership relationship is stable over time. On the counterfactual: equation (A4) invokes conditional independence between missing and observed appliances given income, which is false in a known direction, since appliances co-move on a shared modernisation factor beyond national income, so joint ownership and hence any contribution from positive interaction coefficients is understated; the delta-method standard error propagates uncertainty in $\widehat{\beta}$ only, treating observed and imputed shares as known constants; and averaged baskets need not be internally consistent, in that they can land on a joint-ownership combination no single real country exhibits. To keep these trade-offs auditable, every output row is tagged with the imputation method used, the number and identity of countries that entered only via imputation, the number of missing appliances per country, and the ownership source, so that an imputation-widened estimate is never mistaken for a fully observed one.

B Appendix Tables

Table A1 presents the full three-level activity classification used in the harmonisation of our global time-use database. All time is allocated to one of three mutually exclusive one-digit categories—Work, Education, and Leisure—which are in turn disaggregated into two-digit groups and three-digit activities. By construction, the three-digit codes cover all 1,440 minutes of a day.

Table A1: Activity classification hierarchy (three-digit codes).

Code	Group	Description
1. Active Time		
	Paid work	
111		Household enterprise paid work
112		Outside paid work
	Unpaid work (domestic)	
121		Food preparation & clean-up
122		Cleaning and upkeep
123		Laundry and garment care
124		Production of goods for own use
125		DIY, repair, and domestic services
126		Shopping and budget management
	Unpaid work (care)	
131		Dependent adult care
132		Dependent child care
133		Non-dependent adult care
	Voluntary work	
141		Community voluntary work
142		Voluntary work for other households
150	Production of goods for own use	Production of goods for own use
190	Work commute	Work commute
2. Education		
	Formal education	

continued on next page

Table A1 continued

Code	Group	Description
211		Formal education attendance
212		Homework
	Informal education	
221		Informal educational training
222		Other informal education
290	Education commute	Education commute
3. Leisure		
	Leisure basics	
311		Sleep
312		Eating & drinking
313		Personal care
	Entertainment	
321		Indoors entertainment
322		Outdoors entertainment
	Active leisure	
331		Reading & writing
332		Sports & exercise
333		Games & hobbies
	Social activities	
341		Religious activities
342		Social and community activities
343		Friends & family
390	Leisure commute	Leisure commute

C Appendix Figures

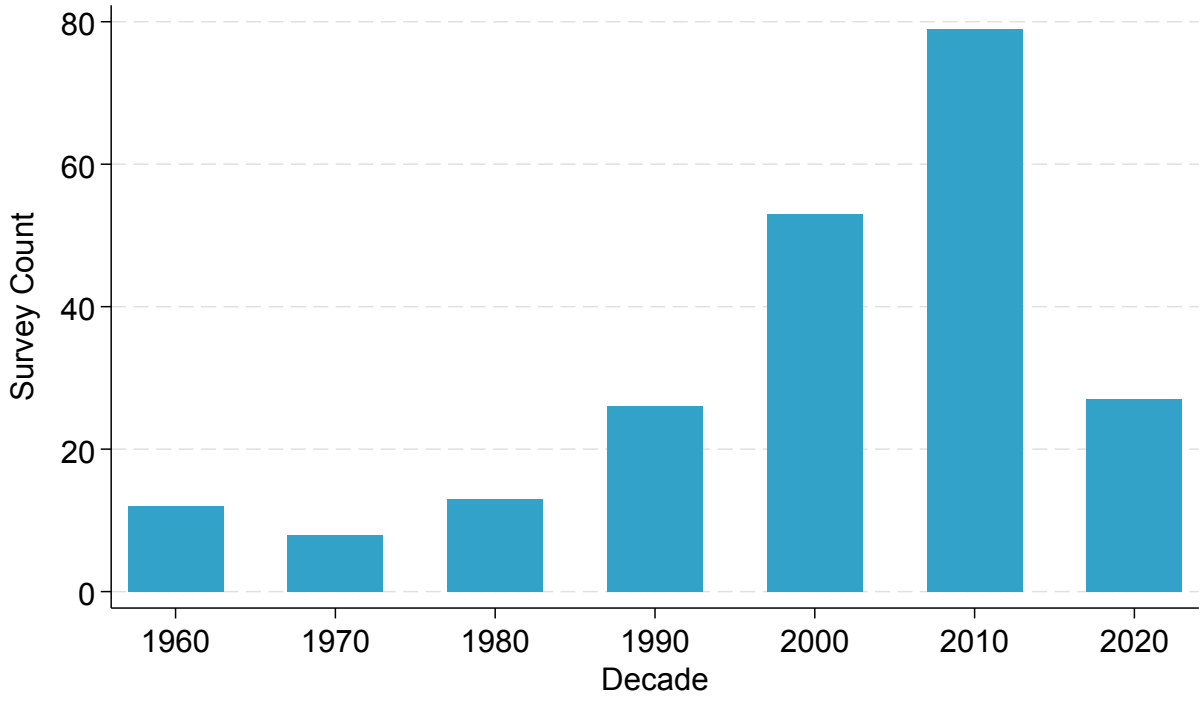


Figure A2: Historical coverage of the time-use database

Notes: Number of time-use surveys in our database by decade. Coverage is sparse before the 1980s, expands steadily through the 1990s and 2000s, and grows rapidly in the 2010s as more countries—particularly in Sub-Saharan Africa and South Asia—began fielding nationally representative time-use surveys for the first time.

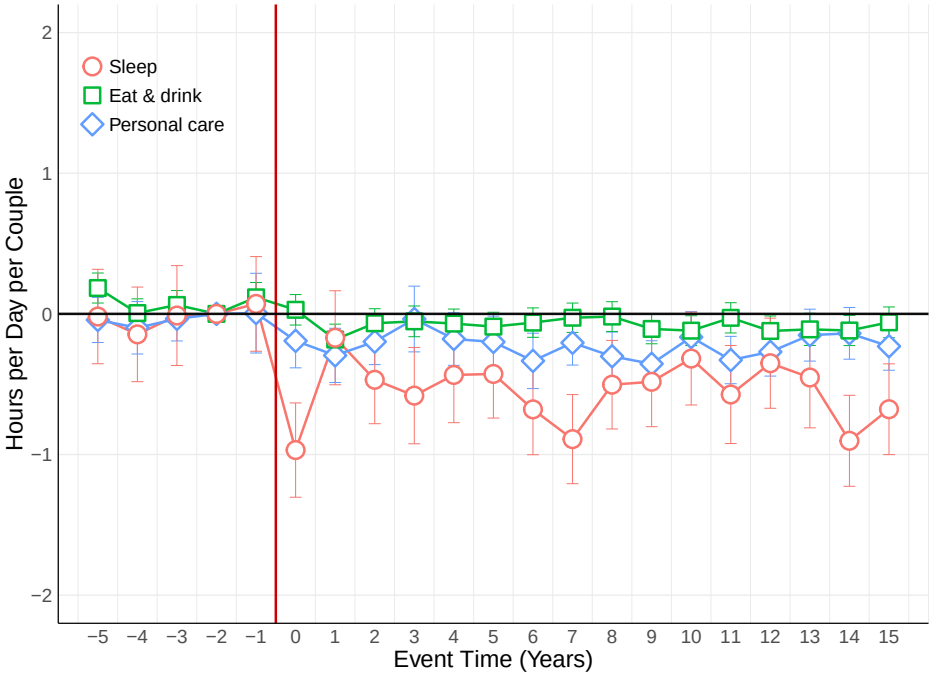


Figure A3: Decomposition of leisure basics

Notes: Couple-level event study around first childbirth decomposing leisure basics into sleep, eating and drinking, and personal care, pooling all countries. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals.

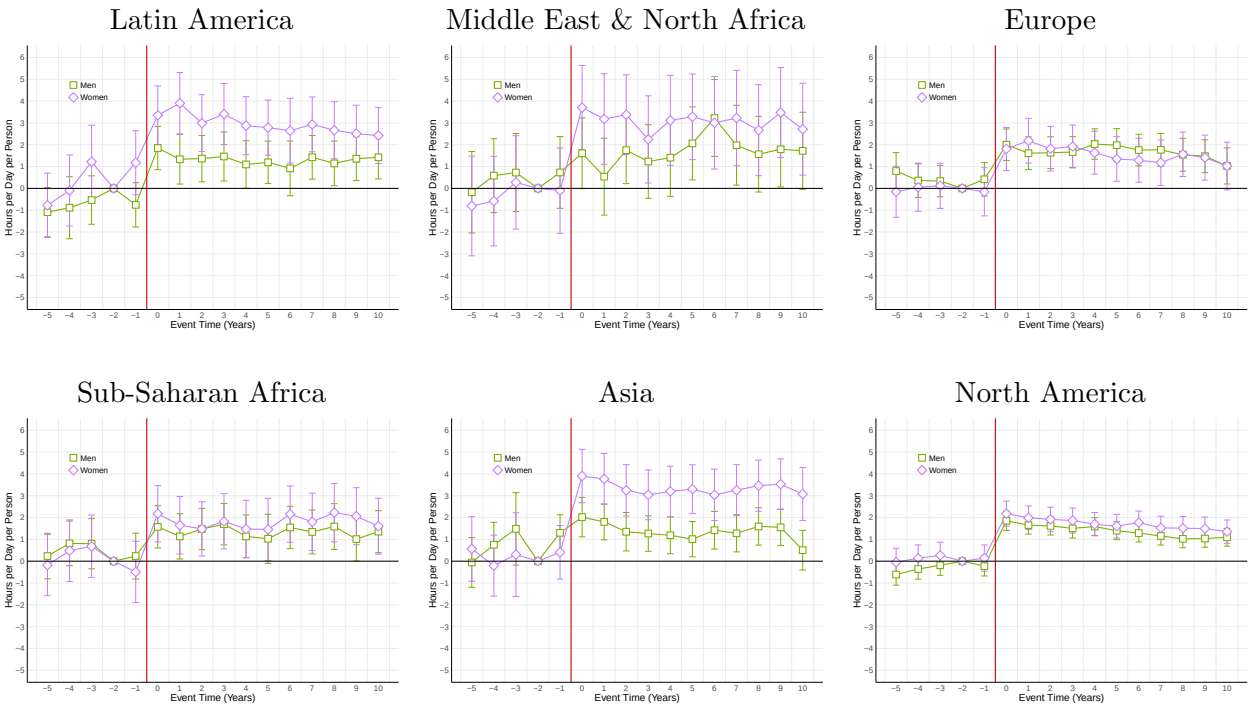


Figure A4: Event Studies for Active time by Region

Notes: Event studies of active time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

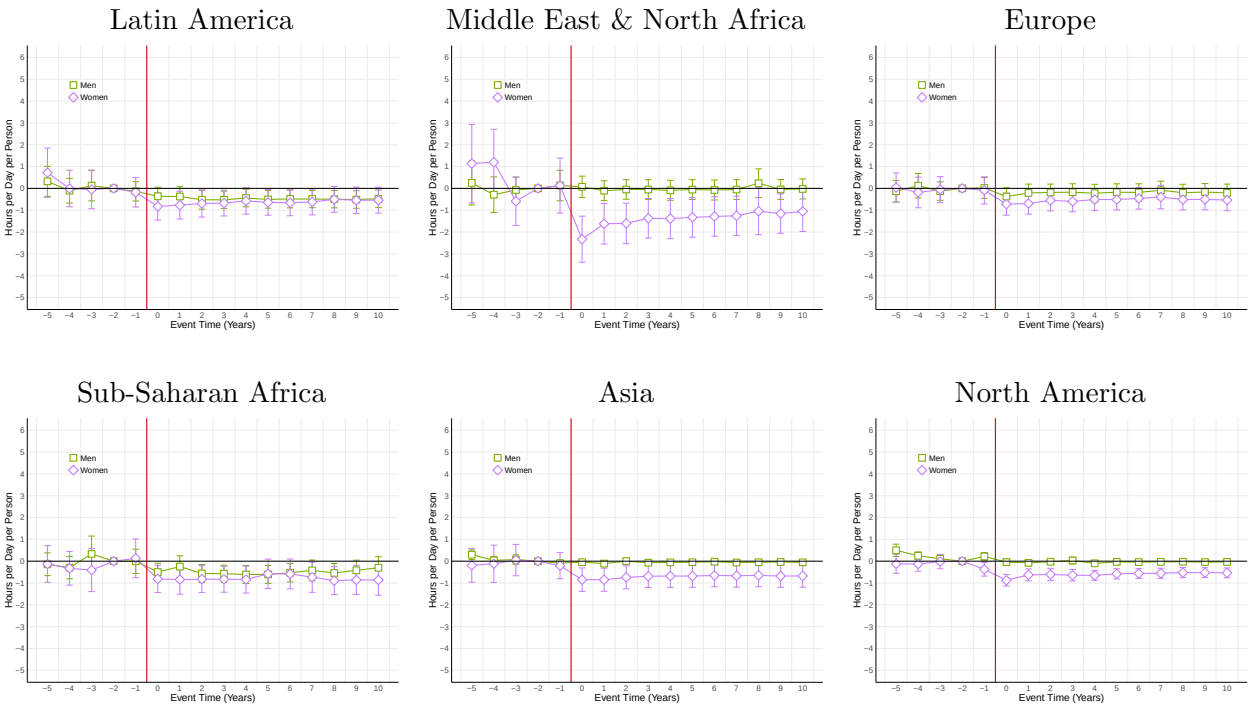


Figure A5: Event Studies for Education by Region

Notes: Event studies of education time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

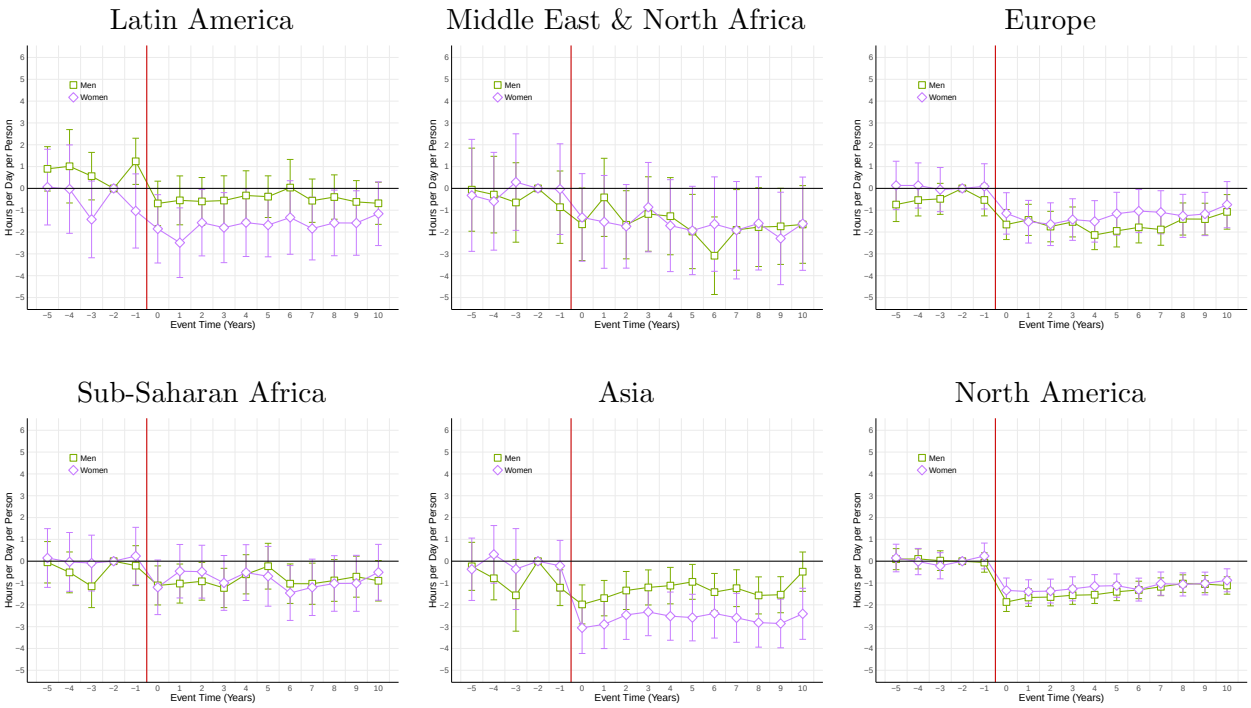


Figure A6: Event Studies for Leisure by Region

Notes: Event studies of leisure time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

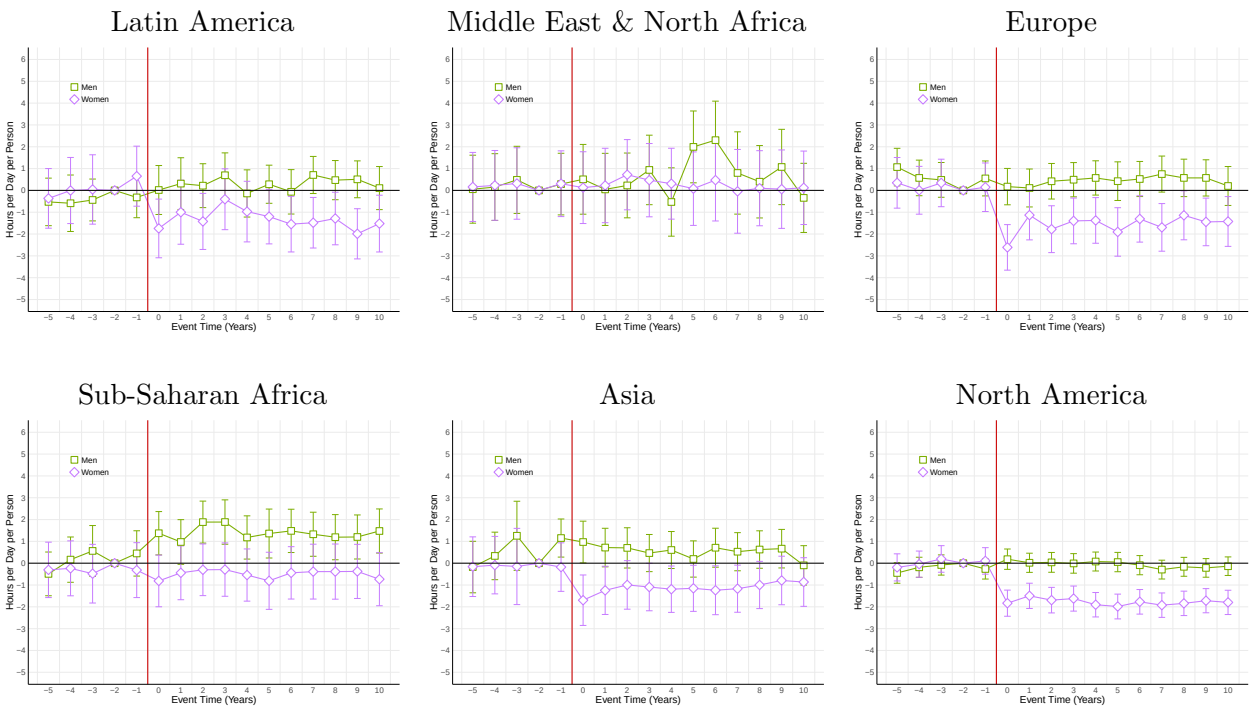


Figure A7: Event Studies for Paid Work by Region

Notes: Event studies of paid work time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

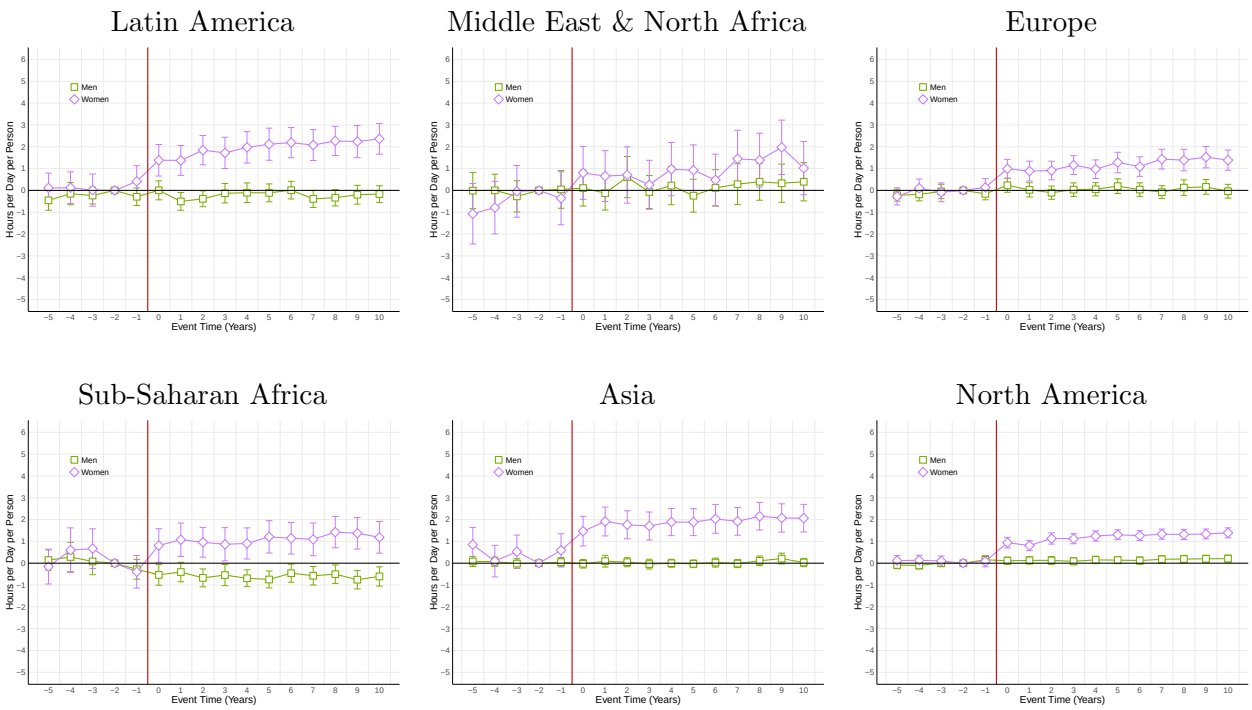


Figure A8: Event Studies for Domestic Work by Region

Notes: Event studies of domestic work time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

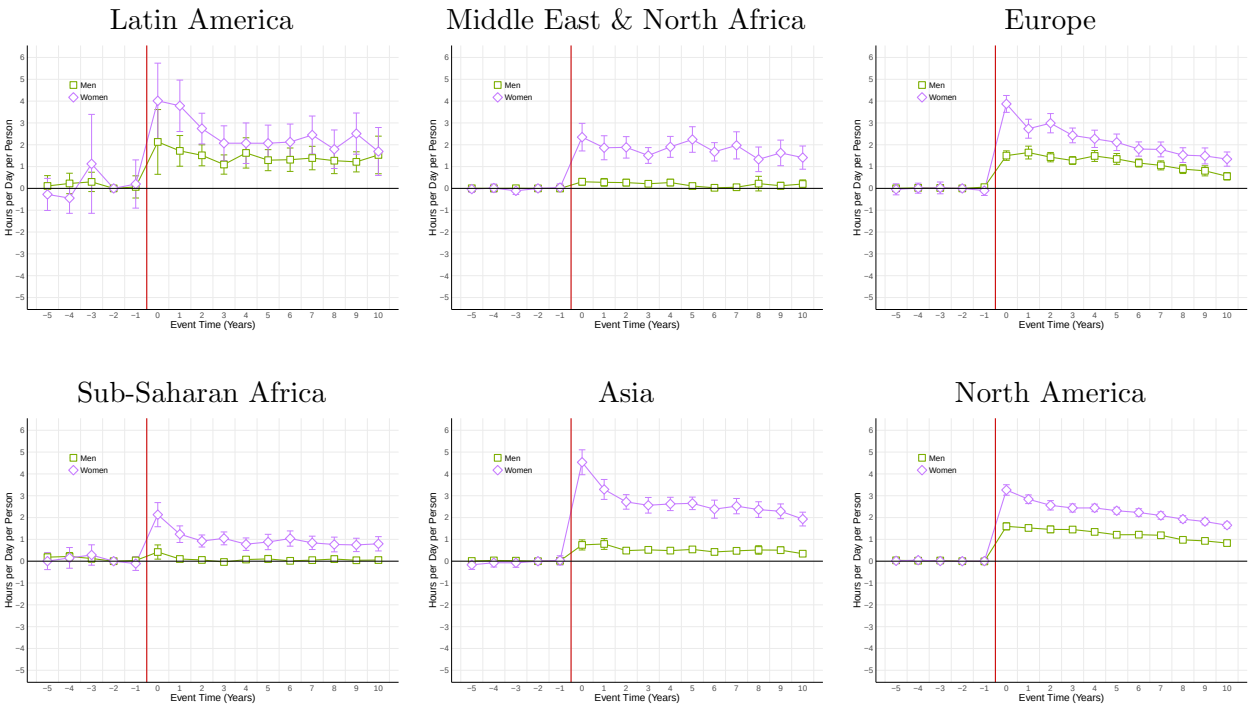


Figure A9: Event Studies for Childcare by Region

Notes: Event studies of childcare time around first childbirth for men and women. The vertical axis shows minutes per day relative to $t = -2$. Shaded bands show 95% confidence intervals. Estimates are pooled across different geographies.

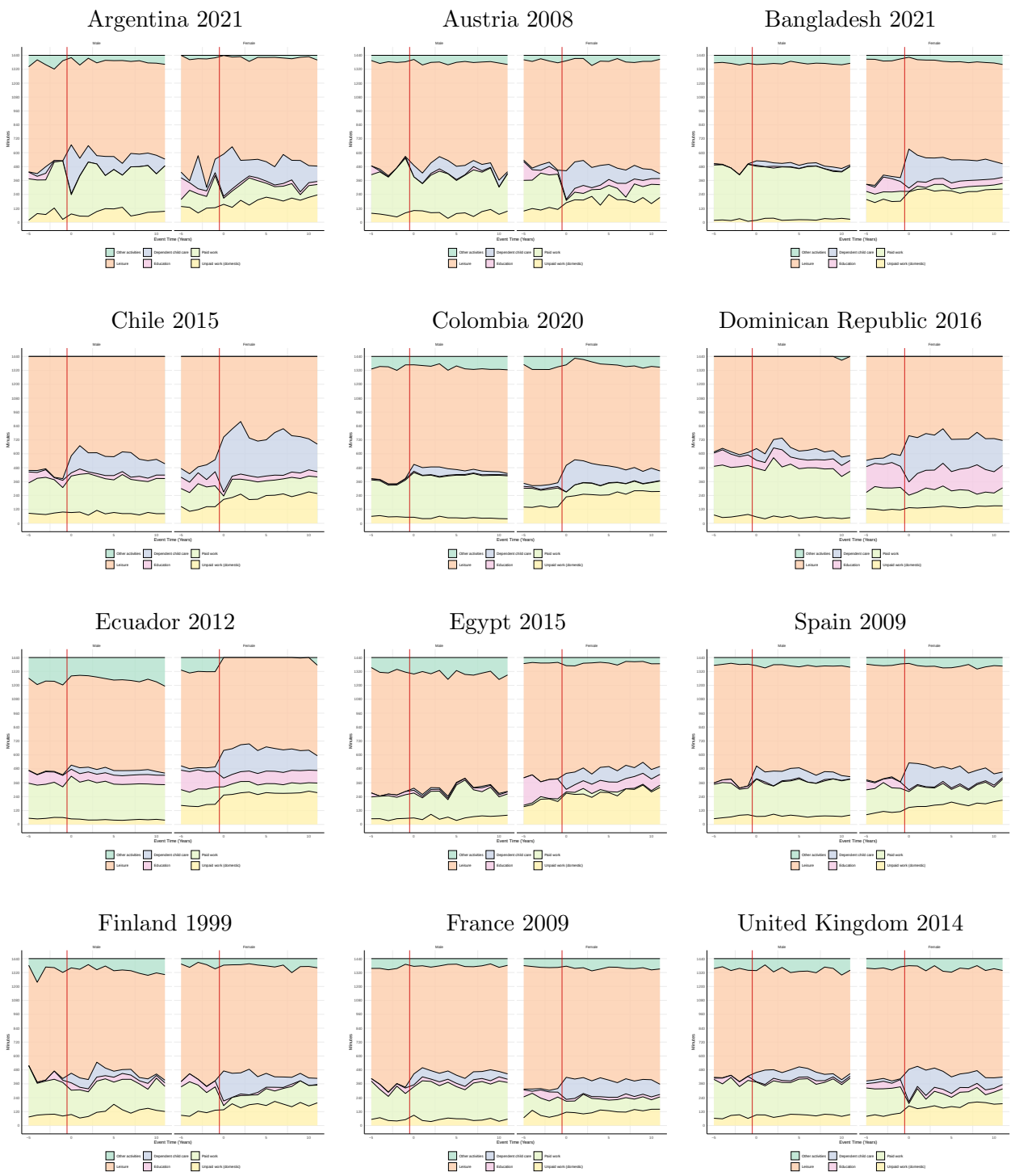


Figure A10: Event Studies of First Childbirth by Country

Notes: Stacked event studies around first childbirth showing the full 24-hour allocation for men (Panel A) and women (Panel B). Each stack shows minutes per day in each time-use category relative to $t = -2$; categories sum to 1,440 minutes per day.

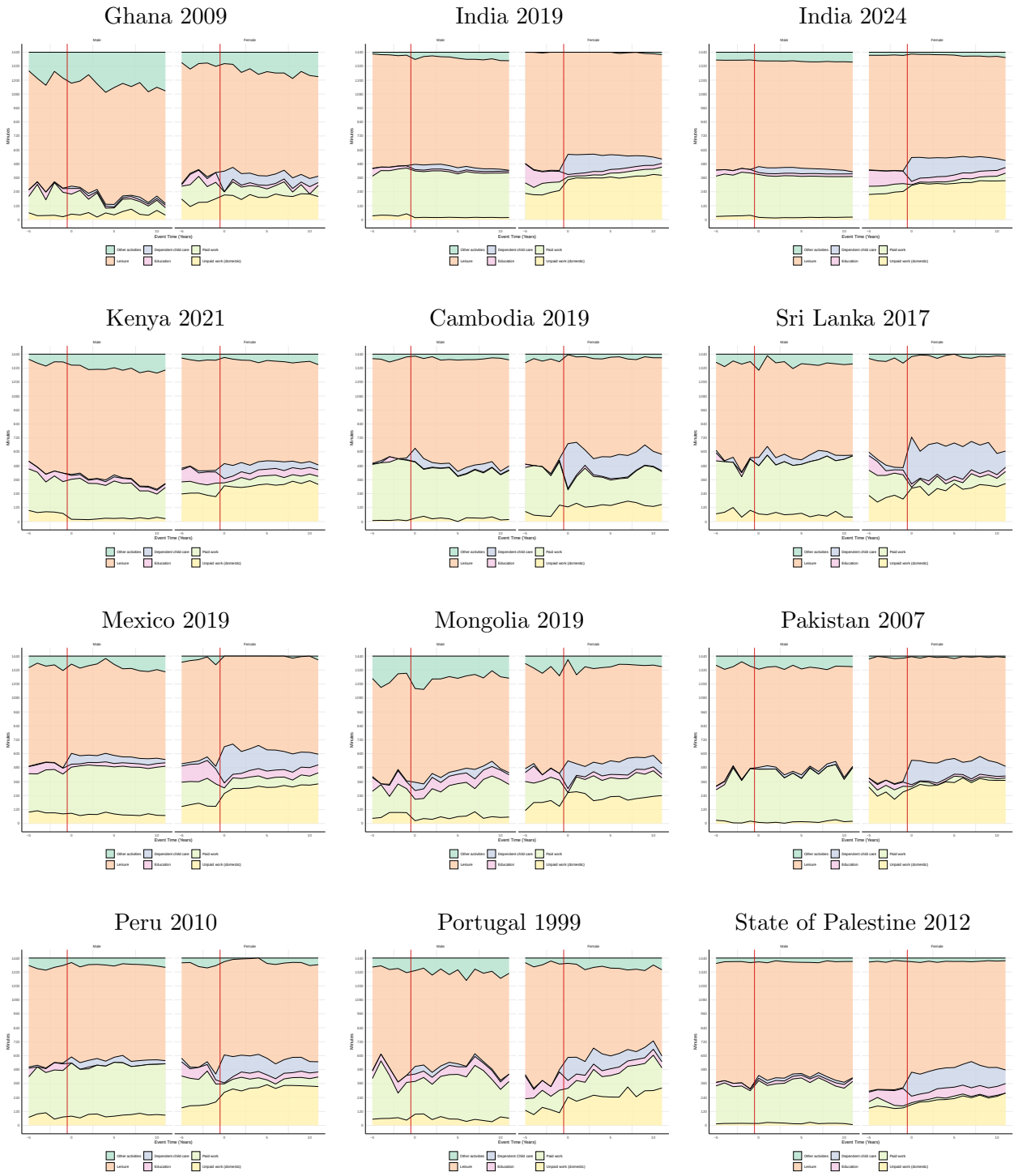


Figure A11: Event Studies of First Childbirth by Country, Continued

Notes: Stacked event studies around first childbirth showing the full 24-hour allocation for men (Panel A) and women (Panel B). Each stack shows minutes per day in each time-use category relative to $t = -2$; categories sum to 1,440 minutes per day.



Figure A12: Event Studies of First Childbirth by Country, Continued

Notes: Stacked event studies around first childbirth showing the full 24-hour allocation for men (Panel A) and women (Panel B). Each stack shows minutes per day in each time-use category relative to $t = -2$; categories sum to 1,440 minutes per day.

D Model Derivations

This appendix provides formal derivations of the first-order conditions and the three quantitative insights discussed in Section 5.3.

D.1 First-order conditions

The household maximises

$$U = W(Q_c, Q_d) + \beta_m v(\ell^m) + \beta_f v(\ell^f) + \beta_y z(y)$$

subject to the time constraints $x_c^k + x_d^k + x_\ell^k + x_w^k = T$ for each spouse $k \in \{m, f\}$ and the budget constraint $w^m x_w^m + w^f x_w^f = y + p_c m_c$. We write $\ell^k \equiv x_\ell^k$ for leisure. Let λ^k denote the shadow price of time for spouse k and μ the multiplier on the budget constraint. The first-order conditions at an interior optimum are:

$$x_w^k: \quad \mu w^k = \lambda^k \tag{A5}$$

$$x_\ell^k: \quad \beta_k (\ell^k)^{-\gamma_k} = \lambda^k \tag{A6}$$

$$x_i^k,\ i \in \{c, d\}: \quad \frac{\partial W}{\partial Q_i}\, \alpha_i^k\, (x_i^k)^{\rho_i - 1}\, Q_i^{1-\rho_i} = \lambda^k \tag{A7}$$

$$m_c: \quad \frac{\partial W}{\partial Q_c}\, \alpha_c^o\, (m_c)^{\rho_c - 1}\, Q_c^{1-\rho_c} = \mu\, p_c \tag{A8}$$

$$y: \quad \beta_y\, y^{-\eta} = \mu \tag{A9}$$

where $\frac{\partial W}{\partial Q_i} = W^{1-\psi}\, \beta_i\, Q_i^{\psi - 1}$, so that (A7) can be written as

$$W^{1-\psi}\, \beta_i\, Q_i^{\psi - \rho_i}\, \alpha_i^k\, (x_i^k)^{\rho_i - 1} = \lambda^k. \tag{A10}$$

Shadow prices. From (A5) and (A9), $\lambda^k = \beta_y\, y^{-\eta}\, w^k$, so:

$$\frac{\lambda^m}{\lambda^f} = \frac{w^m}{w^f}. \tag{A11}$$

Market vs. parental childcare. Dividing (A8) by (A10) for $i = c$, then using (A5) to substitute $\lambda^k = \mu w^k$:

$$\frac{\alpha_c^o\, (m_c)^{\rho_c - 1}}{\alpha_c^k\, (x_c^k)^{\rho_c - 1}} = \frac{\mu\, p_c}{\lambda^k} = \frac{p_c}{w^k}$$

Rearranging and noting $1/(\rho_c - 1) = -\sigma_c$:

$$\frac{m_c}{x_c^k} = \left(\frac{\alpha_c^o}{\alpha_c^k} \cdot \frac{w^k}{p_c} \right)^{\sigma_c}. \tag{A12}$$

Cross-spouse ratio. Writing (A10) for task $i \in \{c, d\}$ for spouse m and spouse f and taking the ratio, the terms $W^{1-\psi}$, β_i, and $Q_i^{\psi-\rho_i}$ cancel; the last equality uses (A5) for each spouse ($\lambda^k = \mu w^k$):

$$\frac{\alpha_i^m \, (x_i^m)^{\rho_i - 1}}{\alpha_i^f \, (x_i^f)^{\rho_i - 1}} = \frac{\lambda^m}{\lambda^f} = \frac{w^m}{w^f}$$

Rearranging:

$$\left(\frac{x_i^m}{x_i^f} \right)^{\rho_i - 1} = \frac{w^m}{w^f} \cdot \frac{\alpha_i^f}{\alpha_i^m}$$

Raising both sides to the power $1/(\rho_i - 1) = -\sigma_i$:

$$\frac{x_i^m}{x_i^f} = \left(\frac{\alpha_i^m}{\alpha_i^f} \cdot \frac{w^f}{w^m} \right)^{\sigma_i}, \quad i \in \{c, d\}. \tag{A13}$$

D.2 Insight 1: Implied norm ratios

Inverting (A13) for $i = c$ and solving for the perceived-productivity ratio:

$$\frac{\alpha_c^f}{\alpha_c^m} = \left(\frac{x_c^f}{x_c^m} \right)^{1/\sigma_c} \cdot \frac{w^f}{w^m}. \tag{A14}$$

Let $R^{\text{data}} \equiv x_c^f / x_c^m$ denote the observed female-to-male childcare ratio and $\omega \equiv w^f / w^m$ the relative female wage. Table A2 reports implied values of α_c^f / α_c^m for a range of empirically relevant parameters.

The table covers the empirically plausible range of σ_c, from strong complementarity ($\sigma_c = 0.25$) to Cobb-Douglas ($\sigma_c = 1$), which serves as the conservative benchmark. Even at $\sigma_c = 1$, the implied norm ratio exceeds 2 for the lowest observed childcare ratios and reaches 5.6 at the high end. For $\sigma_c < 1$ — the case supported by developmental psychology, which documents that mothers and fathers perform qualitatively different types of childcare — the required ratios are substantially larger, reaching double digits at $\sigma_c = 0.75$ and $R^{\text{data}} = 7$. The conclusion that $\alpha_c^f / \alpha_c^m \gg 1$ is robust across the entire range.

Table A2: Implied perceived-productivity ratio α_c^f/α_c^m

	Elasticity of substitution σ_c			
$R^{\text{data}} = x_c^f/x_c^m$	0.25	0.5	0.75	1.0
2.5	31.2	5.0	2.7	2.0
4.0	204.8	12.8	5.1	3.2
7.0	1 920.8	39.2	10.7	5.6

Notes: Each cell reports $\alpha_c^f/\alpha_c^m = (R^{\text{data}})^{1/\sigma_c} \cdot \omega$. The empirical range $R^{\text{data}} \in [2.5, 7]$ is read from Panel (c) of Figure 21. The relative wage $\omega = w^f/w^m$ is set at 0.8, corresponding to the mean raw gender wage gap in low-income countries (19.5%) reported in International Labour Organization (2024), Table 5.1. In middle- and high-income countries the gap is smaller ($\omega \approx 0.87$–0.92), so the implied norm ratios reported here are a lower bound.

D.3 Insight 2: Critical role of β_c

The within-task ratio is invariant to β_c. Equation (A13) shows that the gender ratio of time inputs in task i depends only on perceived productivities α_i^m/α_i^f, relative wages w^f/w^m, and the elasticity σ_i. Preference shifters β_c, β_d, and the complementarity parameter ψ do not appear. Define $R_i \equiv x_i^m/x_i^f < 1$ for $i \in \{c, d\}$ as the (constant) male-to-female ratio in each task:

$$R_i = \left(\frac{\alpha_i^m}{\alpha_i^f} \cdot \frac{w^f}{w^m} \right)^{\sigma_i} < 1. \tag{A15}$$

Condition for childcare to rise despite rising wages. From the expanded FOC (A10) for $i = c$ and spouse f:

$$W^{1-\psi}\, \beta_c\, Q_c^{\psi-\rho_c}\, \alpha_c^f\, (x_c^f)^{\rho_c-1} = \beta_y\, y^{-\eta}\, w^f.$$

Log-differentiate this equation holding W, Q_c, and y fixed (a partial equilibrium that isolates the direct tension between the taste shifter β_c, female childcare time, and the female shadow value of time embedded in w^f). With $\rho_c < 1$,

$$(1 - \rho_c)\, \mathrm{d} \ln x_c^f = \mathrm{d} \ln \beta_c - \mathrm{d} \ln w^f,$$

so $\mathrm{d} \ln x_c^f > 0$ along a path with $\mathrm{d} \ln w^f > 0$ requires $\mathrm{d} \ln \beta_c > \mathrm{d} \ln w^f$. Mapping proportional shifts in an index of development to these log changes yields the benchmark inequality

$$\frac{\mathrm{d} \ln \beta_c}{\mathrm{d} \ln(\text{development})} > \frac{\mathrm{d} \ln w^f}{\mathrm{d} \ln(\text{development})}. \tag{A16}$$

Once W, Q_c, and y adjust endogenously, condition (A16) is not an implication of full comparative statics; it summarises the partial force needed for female childcare time to rise against rising opportunity cost. Since R_c is invariant to β_c, when x_c^f rises we have $x_c^m = R_c\, x_c^f$, so both parents' childcare move together.

A rise in β_c amplifies paid-work divergence. Since R_c is fixed, a rise in β_c that increases x_c^f by $\delta > 0$ increases x_c^m by $R_c\,\delta < \delta$. When $\psi < 1$, Q_c and Q_d are complements in $W(\cdot,\cdot)$, so an increase in Q_c raises $\partial W/\partial Q_d$ and shifts the marginal benefit of domestic time upward in (A10) for $i = d$. In the joint adjustment alongside a β_c-driven expansion of childcare, domestic hours move in the same direction as Q_d; we denote the mother's first-order increment by $\delta_d > 0$ and the father's by $R_d\,\delta_d$, with $R_d < 1$ as in (A15). From the time constraint:

$$dx_w^f = -\delta - \delta_d - d\ell^f \tag{A17}$$

$$dx_w^m = -R_c\,\delta - R_d\,\delta_d - d\ell^m \tag{A18}$$

Since $R_c, R_d < 1$, the reduction in women's paid work exceeds the reduction in men's paid work (for any equal change in leisure). The paid-work gender gap therefore widens:

$$dx_w^m - dx_w^f = (1 - R_c)\,\delta + (1 - R_d)\,\delta_d + (d\ell^f - d\ell^m) > 0 \tag{A19}$$

where the inequality follows from $R_c, R_d < 1$ and the fact that leisure changes are approximately symmetric across spouses. Crucially, when $R_c = R_d = 1$, the first two terms vanish: β_c raises home production equally for both spouses and paid-work divergence from the home-production margins disappears. This formalises the claim that $\beta_c \uparrow$ widens the paid-work gender gap in the direction observed in the data when $R_c < 1$ and $R_d < 1$, as in our calibration, and not when both ratios equal one so that marginal home production expands in lockstep across spouses.

D.4 Insight 3: Women's return to paid work requires a domestic-work revolution

Time-constraint accounting. From the female time constraint:

$$x_w^f = T - x_c^f - x_d^f - \ell^f. \tag{A20}$$

Over the development path we observe (i) x_c^f rising monotonically (Figure 21, Panel (c) and Insight 2), (ii) ℓ^f approximately constant, and (iii) x_w^f increasing at high levels of development (Figure 21, Panel (a)). Differentiating:

$$\underbrace{dx_w^f}_{>0} = -\underbrace{dx_c^f}_{>0} - dx_d^f - \underbrace{d\ell^f}_{\approx 0}. \tag{A21}$$

For $dx_w^f > 0$ to hold despite $dx_c^f > 0$, it must be that $dx_d^f < -dx_c^f$. In words, mothers' domestic work must fall by more than the increase in their childcare.

Cheaper market childcare cannot close the gap. One candidate mechanism is a fall in the price of market childcare p_c. From (A12), a decline in p_c raises the market-to-parental care ratio m_c/x_c^k with elasticity σ_c. When σ_c is low (parental and market childcare are complements), this substitution is limited: cheaper nurseries or preschools cannot easily replace parental time. Moreover, the cross-spouse ratio (A13) is invariant to p_c, so a decline in the price of market care reduces both parents' childcare proportionally and does not free up maternal time differentially. Empirically, the fact that total parental childcare *rises* throughout the development path is consistent with complementarity between parental and market care inputs.

A productivity revolution in domestic work. To allow for technological change in domestic work, augment the production function with a Hicks-neutral productivity parameter A_d:

$$Q_d = A_d \left[\alpha_d^m (x_d^m)^{\rho_d} + \alpha_d^f (x_d^f)^{\rho_d}\right]^{1/\rho_d}. \tag{A22}$$

An increase in A_d — representing the introduction of household appliances, market substitutes for cooking and cleaning, or improvements in domestic infrastructure — allows the household to maintain Q_d at a given level while reducing time inputs. Letting $Z_d \equiv [\alpha_d^m (x_d^m)^{\rho_d} + \alpha_d^f (x_d^f)^{\rho_d}]^{1/\rho_d}$ so that $Q_d = A_d Z_d$, the FOC for x_d^f becomes:

$$W^{1-\psi}\, \beta_d\, Q_d^{\psi-\rho_d}\, A_d^{\rho_d}\, \alpha_d^f\, (x_d^f)^{\rho_d - 1} = \lambda^f.$$

An increase in A_d raises the left-hand side both directly (through $A_d^{\rho_d}$) and indirectly (through $Q_d^{\psi-\rho_d}$), increasing the effective marginal product of time in domestic work. To restore equality, x_d^f must fall. Through the time constraint (A20), this freed-up time can be reallocated to paid work.

The cross-spouse ratio for domestic work remains governed by (A13) regardless of A_d, so

the productivity gain reduces *both* spouses' domestic work proportionally. However, since $x_d^f \gg x_d^m$ (due to norms $\alpha_d^f/\alpha_d^m \gg 1$), the absolute time savings are much larger for mothers. This is the domestic-work revolution: the mechanism through which rising female labour force participation at high income is achieved, not through a decline in childcare, but through a decline in the time cost of domestic work.