Policy Research Working Paper 10305

Maternal Work and Children's Development

Examining 20 Years of Evidence

Maria C. Lo Bue
Elizaveta Perova
Sarah Reynolds

WORLD BANK GROUP
East Asia and the Pacific Region
Office of the Chief Economist
February 2023

Policy Research Working Paper 10305

Abstract

Maternal work may affect children positively through increased household income, higher control of mothers over available income, and expansion of maternal information networks through work contacts and greater decision-making power of mothers as they become more economically empowered. However, maternal work may reduce maternal time spent with children. If maternal time is not substituted for time of equal quality by other caregivers, children's development may be penalized. Stress associated with work may also decrease the quality of parenting. This review summarizes causal evidence on the relationship between maternal work and children's development. The majority causal studies find positive or null impacts of maternal work on children's development.

This paper is a product of the Office of the Chief Economist, East Asia and the Pacific Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at eperova@worldbank.org.

Maternal Work and Children's Development: Examining 20 Years of Evidence

Maria C. Lo Bue, Elizaveta Perova and Sarah Reynolds

J13, J21, J16, E24

Keywords: child development, maternal labor force participation, gender equality, review of evidence

Complex links between maternal work and children's development

Maternal work is linked to children's development through several channels. On the one hand, maternal work may affect children's development positively. First, children benefit from living in households with higher income, and mothers—in comparison to fathers—spend a higher proportion of their resources on children (Lundberg 1997; Ermish and Francesconi 2013; Dunbar, Dunbar, Lewbel, and Pendakur, 2013). There are also intangible benefits of maternal work, such as mothers' exposure to a larger social network through work contacts, greater access to information, and a potential increase in mothers' decision-making power, all of which may allow for better choices around monetary and non-monetary investments in children (Currie, 2009; Doss, 2013; Lépine and Strobl, 2013).

On the other hand, maternal work may affect children's development negatively through reduction in maternal time spent with children and, consequently, fewer opportunities for interactions and consistent engagement. Child development would be penalized by maternal work if institutional or other non-maternal childcare is of inferior quality than maternal care and the increase in income does not compensate with nutritional and safety expenses, for example. Stress associated with work may reduce the quality of parenting (Baker, Gruber, Milligan, 2008). The combination of these factors may lead to less secure attachment with the mother - an important factor for children's socio-emotional development and subsequent academic success (Alto and Petrenko, 2017; Allen, 2008). Maternal work may also increase the likelihood that older children take on more domestic responsibilities and spend less time on homework or enrichment activities (Afridi et al., 2016).

With both positive and negative theoretical impacts of mothers' work on child development, it is important to examine whether up-to-date empirical evidence suggests overall positive, negative, or null impacts. Our study summarizes the evidence of maternal work on child and adolescent health and intellectual development from the past 20 years. We note strengths and limitations of the studies and highlight common threads relevant to different facets of policy design, including considerations for emerging economies.

Broad coverage of contexts and measures

Using a variety of search terms, we retreived 1,181 studies from databases that cover medical, psychological, economic and other social science literature. We identified a total of 613 relevantly-titled articles published between 2002 and 2021, and further limited them to 312 which examined the links between maternal work and development of children ages 0 to 18 years. We further narrowed our selection to 80 papers, which utilized statistical methods to plausibly establish causal impact. Within these studies, we limited our analysis to a set of 26 articles, which used the most rigourous methods: instrumental variables, sibling fixed effects, and multiple analytic approaches (most commonly individual fixed effects and lagged dependent variables combined).

Twenty-six causal papers cover a range of contexts: 21 papers are from Western countries (Australia, Denmark, Italy, Norway, the United Kingdom, and the United States) and 5 are from non-Western: three lower-middle income (the Arab Republic of Egypt, India, and Indonesia) and one high income (Chile). Among papers on high-income countries (all Western countries and Chile), 5 focus on low-income populations (such as mothers of welfare in the United States), and 22 rely on nationally representative surveys or censuses.

Overall, the 26 causal studies examined 20 different measures of child development. To reduce multiple dimensions of this body of evidence, we combined them into 6 main outcomes: schooling, cognitive, behavioral and 3 types of health measures. Schooling outcomes included years of schooling, current enrollment, grade attaintment and high school grades. Cognitive outcomes included test scores from a range of psycometrically validated exams. Behavioral outcomes are largely comprised of self-reported risky beahviors for adolescents and parent-reported standardized indexes for behavioral problems for younger children. We grouped health outcomes into three categories: (1) height outcomes, which capture long-term nutrition, (2) weight outcomes, which consider short-term nutrition and physical activity,[1] and (3) other health measures, most of which were indicators of medical concerns such as lung capacity, hemoglobin levels, asthma episodes, ear infections and overnight hospitalization.

Several studies examined more than one group of outcomes. To facilitate the discussion of results, we refer to separate groups of outcomes within a study as a substudy. A study can have one substudy or multilple substudies. Twenty-six causal studies yielded 40 substudies with evidence on 6 outcomes (height, weight, health, schooling, behavior and cognition).

For each substudy, we categorized whether it detected null, positive or negative impacts of maternal work on child development for each group of outcomes it examined. If at least one outcome in a group is negative/positive in a statistically significant way, while impacts on other outcomes in that group are not statistically significant, we count it as evidence of negative/positive impact for the substudy.[2] If there was no statistically significant impact on any outcomes in the group, we count it as evidence of null effect. There were no papers which detected a statistically significant positive impact on one outcome in a group, and negative on a different one.

Main findings: Impacts on children's development

The most abundant evidence is on weight-related outcomes, followed by cognitive and behavioral outcomes (12, 8 and 7 substudies, respectively). The evidence is more scant on schooling, health and height outcomes, with 5, 4 and 4 substudies, respectively.

The majority of causal evidence from 10 countries over the last 20 years suggests that maternal work does not hurt children's development. Of the substudies, 68% (27 out of 40) do not find evidence of negative impact, with 6 finding positive, and 21 null effects (Figure 1).

An interesting feature of the studies reviewed is that many of them use a wide range of outcomes, and only 2 corrected for multiple hypotheses testing, or adjusted for the probability of discovering a non-null effect purely by chance. While common practice in medicine, correction for multiple hypotheses testing started making forays into social science only recently.[3] Thus the 20-year span of our analysis incorporates many papers that were not yet submitted to this additional check.

[1] Weight outcomes include deviations from the norm in both directions: underweight as well as obesity.

[2] For example, the study in Indonesia examines 2 indicators for health: hemoglogin and lung capaicty. The authors find statistically signficant impacts on hemoglobin, but not on lung capacity. We categorize this result as evidence of positive impact on health.

[3] Some of the earliest papers that raised the question of correcting for multiple hypotheses testing and demonstrating sensitivity of results include Anderson (2008), List, Shaikh and Xu (2019), Kling, Liebman and Katz (2007).

Using the values of coefficients and standard errors in the papers, we have carried out two types of checks for multiple hypotheses testing, using both Bonferroni correction and the less conservative Benjamini-Hochberg (1995) correction for False Discovery Rate. Application of these corrections reduces the number of substudies with negative results to 9,[4] with 78% of the substudies yielding non-negative results (Figure 2).

Understanding when maternal work may have a negative impact

If one of the objectives of the policy maker is to do no harm, it is important to focus on the negative sub-studies and understand whether there are specific features of maternal work that may have negative impact on children. A closer examination of papers with negative outcomes reveals several trends.

First, the environment in which children of working mothers are while mothers are working matters. Both Rashad and Sharaf (2019) and Shajan and Subbyamoola (2020) find that maternal work increases the likelihood that children are stunted in Egypt and India, respectively. These two countries have low presence of early childhood education facilities, compared to other countries analyzed in the papers reviewed. Thus, the children of working mothers are likely to be with them in the field or in the market, rather than in center-based care. Indeed, Shajan and Subbyamoola (2020) point out that their results are driven by children of women employed in agricultural and manual, rather than professional jobs. Gennetian et al. (2010) examine the impacts of maternal work for low-income populations in the United States, focusing on women eligible for the public welfare program. The earnings from the public welfare program may not be sufficient to cover the costs of high-quality childcare.

Second, specific aspects of maternal work matter. Felfe and Hsin (2012) examine the impacts of characteristics of maternal work, comparing working mothers exposed to more or less hazardous environments. Thus, their study suggests the negative impact of work-related stress and hazards, rather than maternal work per se. James-Burdumy (2005) finds that maternal work results in negative cognitive impacts at age 9 only during the first year of life. However, work during the second and third years of life does not have this effect, suggesting that children's vulnerability or sensitivity to mothers' working is in very early childhood. Additionally, the cognitive impact is not all-around: only math scores were impacted, while reading and vocabulary scores had positive, albeit not significant association with maternal work.

Third, in developed countries, when viable childcare options are present, the impacts are generally low. James-Burdumy (2005) notes that the size of the "loss" to mathematics skill is relatively small at a 0.03 effect size. Furre Haaland, Rege and Votruba (2013) find that 5 additional years of full-time employment reduce a child's education by 0.065 year (or 4 percent of the standard deviation in their sample).

Lastly, one study pertains to a relatively narrow context, which allows for a potentially different interpretation of outcomes. Morrill (2011) focuses on the impact on hospitalizations in the United States. Notably, her analysis does not control for whether the mother has insurance. In the United States, the availability of insurance may be an important factor in the decision whether to take the child to the doctor or not, conditional on the same level of severity of incident. Mothers who work are more likely to have insurance. Thus, the paper may be capturing a different treatment for the same level of severity of

[4] The sub-studies with non-null results which did not withstand correction for multiple hypotheses testing found negative behavioral impact (Felfe and Hsin, 2012 and Aughinbaugh and Gittleman, 2003); negative impact on being underweight (Rashad and Sharaf, 2012), and increase in the likelihood of obesity (Von Hinke Kessler Scholder, 2008).

symptoms, and plausibly, higher likelihood to seek medical attention may be considered a positive outcome.

Discussion

Of the 90 countries where World Values Surveys were administered between 2017 and 2022, in about half of the countries half of the respondents either agree or strongly agree that a preschool child suffers with a working mother. The rate of agreement ranges from 9% in Denmark to 88% in Bangladesh. In three-quarters of the countries, at least 30% of respondents agree that maternal work before primary school is detrimental for the child (Haerpfer et al., 2022).

However, the rigorous research over the past 20 years does not support this widely shared opinion. Most of the papers find non-negative impact of maternal work on children. Several studies that do find negative impacts suggest that specific aspects of maternal work, such as stress or timing, or circumstances in which women work (such as lack of availalbity of high-quality childcare) may be driving negative impacts, rather than maternal work per se. Fortunately, these parameters can be addressed by policy: through improving working conditions, supporting provision of childcare and offering parental leave.

Maternal work is important for economic growth and welfare. Women's entry into the labor force has been a cornerstone of economic growth during the last 50 years in the United States (Hsieh et al., 2019). Currently, a number of governments (including the United States and France) are increasing their budgetary allocations to childcare services in order to support maternal work. Our review does not point to inevitable hidden costs of maternal work in the form of human capital losses of children; rather, it points to the necesity to shape the conditions in which women work through parental leave and childcare policies. Such policy measures are likely to ensure that the economies can reap the short-term economic benefits of maternal work without long-term losses in human capital.

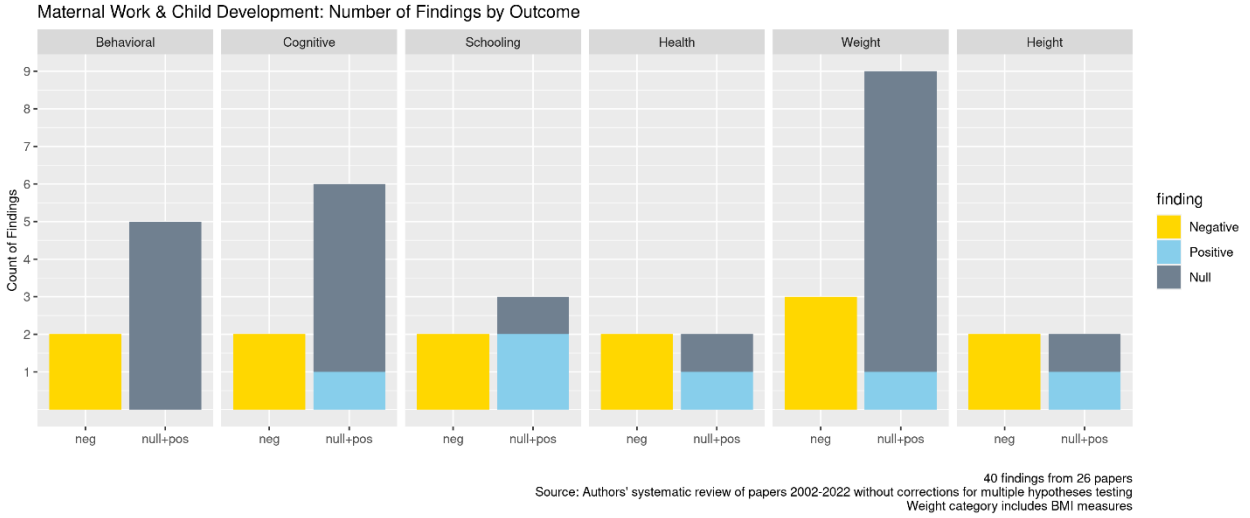


Figure 1

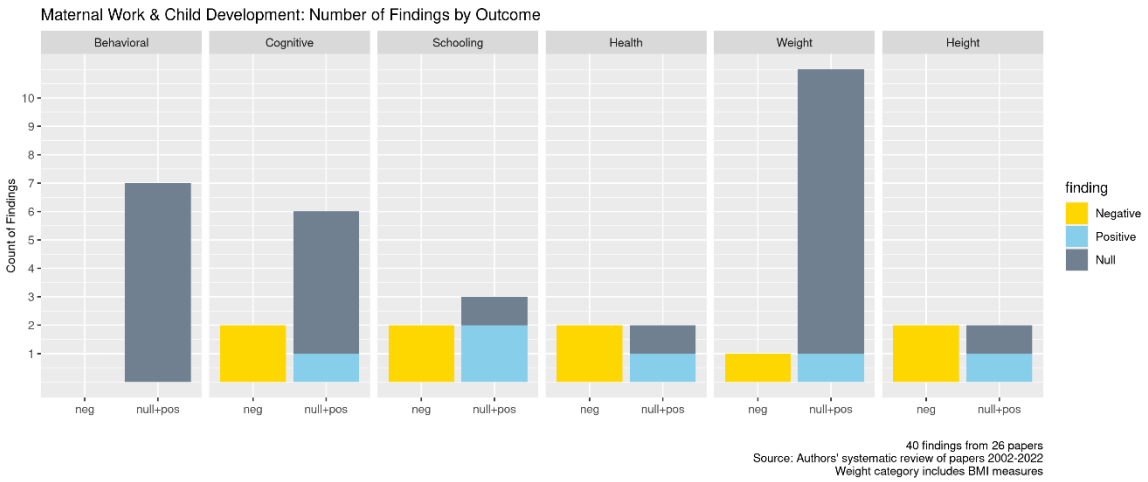

Maternal Work & Child Development: Number of Findings by Outcome
Behavioral
Cognitive
Schooling
Health
Weight
Height
Count of Findings
10
9
8
7
6
5
4
3
2
1
neg
null+pos
finding
Negative
Positive
Null
40 findings from 26 papers
Source: Authors' systematic review of papers 2002-2022
Weight category includes BMI measures

Figure 2

References

Afridi, F., Mukhopadhyay, A. & Sahoo, S. Female labor force participation and child education in India: evidence from the National Rural Employment Guarantee Scheme. IZA J Labor Develop 5, 7 (2016). https://doi.org/10.1186/s40175-016-0053-y

Allen, Joseph P. 2008. "The Attachment System in Adolescence." In Handbook of Attachment: Theory, Research, and Clinical Applications, 2nd Ed, 419–35. New York, NY, US: The Guilford Press.
Alto, Michelle E., and Christie L. M. Petrenko. 2017. "Fostering Secure Attachment in Low- and Middle-Income Countries: Suggestions for Evidence-Based Interventions." Evaluation and Program Planning 60 (February): 151–65. https://doi.org/10.1016/j.evalprogplan.2016.11.003.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." Journal of Political Economy 116, no. 4 (August 2008): 709–45. https://doi.org/10.1086/591908.

Currie, J. (2009). 'Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development'. Journal of Economic Literature, 47(1): 87–122. https://doi.org/10.1257/jel.47.1.87

Cheryl Doss, Intrahousehold Bargaining and Resource Allocation in Developing Countries, The World Bank Research Observer, Volume 28, Issue 1, February 2013, Pages 52–78, https://doi.org/10.1093/wbro/lkt001

Duflo, Esther. 2003. "Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa." World Bank Economic Review 17 (1): 1–25.

Dunbar, Geoffrey R, Arthur Lewbel, and Krishna Pendakur. "Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi." American Economic Review 103, no. 1 (February 1, 2013): 438–71. https://doi.org/10.1257/aer.103.1.438.

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin & B. Puranen (eds.). 2022. World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi:10.14281/18241.20

Kling, J., Liebman, J., and Katz, L. (2007). Experimental analysis of neighborhood effects. Econometrica, 75(1):83–119.

Lépine, A., and E. Strobl (2013). 'The Effect of Women's Bargaining Power on Child Nutrition in Rural Senegal'. World Development, 45: 17–30. https://doi.org/10.1016/j.worlddev.2012.12.018

John A. List & Azeem M. Shaikh & Yang Xu, 2019. "Multiple hypothesis testing in experimental economics," Experimental Economics, Springer;Economic Science Association, vol. 22(4), pages 773-793, December.

Lundberg, S. J, R. A Pollak, and T. J Wales. 1997. "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit." *Journal of Human Resources* 32 (3): 463–80.

Reviewed Papers

Afridi, F., Mukhopadhyay, A., & Sahoo, S. (2016). Female labor force participation and child education in India: Evidence from the National Rural Employment Guarantee Scheme. *IZA Journal of Labor & Development*, *5*(1), 7. https://doi.org/10.1186/s40175-016-0053-y

Agiro, B. T., & Huang, W.-C. (2020). Re-Examining the Effect of Maternal Employment on Child Overweight: The Case of School-Age Children. *Journal of Family and Economic Issues*, *41*(1), 140–157. https://doi.org/10.1007/s10834-019-09658-8

Anderson, P. M., Butcher, K. F., & Levine, P. B. (2003). Maternal employment and overweight children. *Journal of Health Economics*, *22*(3), 477–504. https://doi.org/10.1016/S0167-6296(03)00022-5

Aughinbaugh, A., & Gittleman, M. (2004). Maternal employment and adolescent risky behavior. *Journal of Health Economics*, *23*(4), 815–838. https://doi.org/10.1016/j.jhealeco.2003.11.005

Baum, C. L. (2015). Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking. *Journal of Labor Economics*. https://doi.org/10.1086/345563

Bishop, J. (2011). The Effect of Maternal Employment on Youth Overweight in Australia. *Economic Record*, *87*(s1), 92–104. https://doi.org/10.1111/j.1475-4932.2011.00747.x

Del Boca, D., Pasqua, S., & Suardi, S. (2016). Child Care, Maternal Employment, and Children's School Outcomes. An Analysis of Italian Data. *European Journal of Population*, *32*(2), 211–229. https://doi.org/10.1007/s10680-015-9370-0

Dervisevic, E., Perova, E., & Lo Bue, M. C. (2021). *Maternal employment and children's outcomes: Evidence from Indonesia*. The World Bank.

Dunifon, R., Kalil, A., Crosby, D., & Su, J. H. (2013). Mothers' Night Work and Children's Behavior Problems. *Developmental Psychology*, *49*(10), 1874–1885. https://doi.org/10.1037/a0031241

Ermisch, J., & Francesconi, M. (2013). The Effect of Parental Employment on Child Schooling. *Journal of Applied Econometrics*, *28*(5), 796–822. https://doi.org/10.1002/jae.2260

Felfe, C., & Hsin, A. (2012). Maternal work conditions and child development. *Economics of Education Review*, *31*(6), 1037–1057. https://doi.org/10.1016/j.econedurev.2012.06.008

Gennetian, L. A., Hill, H. D., London, A. S., & Lopoo, L. M. (2010). Maternal employment and the health of low-income young children. *Journal of Health Economics*, *29*(3), 353–363. https://doi.org/10.1016/j.jhealeco.2010.02.007

Greve, J. (2011). New results on the effect of maternal work hours on children's overweight status: Does the quality of child care matter? *Labour Economics*, *18*(5), 579–590. https://doi.org/10.1016/j.labeco.2011.03.003

Haaland, V. F., Rege, M., & Votruba, M. (2013). *Nobody Home: The Effect of Maternal Labor Force Participation on Long-Term Child Outcomes* (SSRN Scholarly Paper ID 2363912). Social Science Research Network. https://doi.org/10.2139/ssrn.2363912

Hubbard, M. N. (2009). *Effects of maternal employment and child care on the health of young children* [Ph.D., The University of North Carolina at Chapel Hill]. https://www.proquest.com/docview/304960202/abstract/4C8AF9FF2D534249PQ/1

James-Burdumy, S. (2005). The Effect of Maternal Labor Force Participation on Child Development. *Journal of Labor Economics*, *23*(1), 177–211. https://doi.org/10.1086/425437

Mendolia, S. (2016). Maternal Working Hours and the Well-Being of Adolescent Children: Evidence from British Data. *Journal of Family and Economic Issues*, *37*(4), 566–580. https://doi.org/10.1007/s10834-015-9480-1

Morrill, M. S. (2011). The effects of maternal employment on the health of school-age children. *Journal of Health Economics*, *30*(2), 240–257. https://doi.org/10.1016/j.jhealeco.2011.01.001

Pekkurnaz, D. (2014). *Dynamic analysis of maternal employment, child care, quality and early childhood development* [Ph.D., The University of North Carolina at Chapel Hill]. https://www.proquest.com/docview/1613182648/abstract/D6F7129E26342F5PQ/1

Pilkauskas, N. V., Brooks-Gunn, J., & Waldfogel, J. (2018). Maternal Employment Stability in Early Childhood: Links with Child Behavior and Cognitive Skills. *Developmental Psychology*, *54*(3), 410–427. https://doi.org/10.1037/dev0000438

Rashad, A. S., & Sharaf, M. F. (2019). Does maternal employment affect child nutrition status? New evidence from Egypt. *Oxford Development Studies*, *47*(1), 48–62. https://doi.org/10.1080/13600818.2018.1497589

Reynolds, S. A., Fernald, L. C. H., & Behrman, J. R. (2017). Mothers' labor market choices and child development outcomes in Chile. *Social Science and Medicine - Population Health*, *3*, 756–766. https://doi.org/10.1016/j.ssmph.2017.08.010

Shajan, J., & Sumalatha, B. S. (2020). Maternal employment and children's health in India: An exploratory analysis. *Journal of Public Affairs*, *n/a*(n/a), e2580. https://doi.org/10.1002/pa.2580

von Hinke Kessler Scholder, S. (2008). Maternal employment and overweight children: Does timing matter? *Health Economics*, *17*(8), 889–906. https://doi.org/10.1002/hec.1357

Waldfogel, J., Han, W.-J., & Brooks-Gunn, J. (2002). The effects of early maternal employment on child cognitive development. *Demography (Pre-2011)*, *39*(2), 369–392.

Ziol-Guest, K. M., Dunifon, R. E., & Kalil, A. (2013). Parental employment and children's body weight: Mothers, others, and mechanisms. *Social Science & Medicine*, *95*, 52–59. https://doi.org/10.1016/j.socscimed.2012.09.004

Supplementary Materials for

Maternal Work and Children's Development: Examining 20 Years of Evidence

By Maria C. Lo Bue, Elisaveta Perova and Sarah Reynolds

1. Methods

1.2 Search

To identify previous studies, we first listed search terms determined from an informal scoping review: we considered the main types of outcomes researched and used the authors' knowledge of common realms of outcomes in the child development. (All authors had published on this topic previously.) The following three sets of search terms emerged: "Maternal" OR "Mother"; "Work" OR "employment" OR "labor"; and "child development" OR "health" OR "nutrition" OR "obesity" OR "stunting" OR "wasting" OR "underweight" OR "overweight" OR "behavior" OR "socio-emotional" OR "cognition" OR "schooling" OR "education" OR "illness" OR "asthma accident".

In October 2021 a research assistant used the 104 search term permutations to collect citations from EconLit and Web of Science.[5] From this list, we selected the top three permutations with the most results and searched them in PubMed. This search did not provide additional documents to include in the review, so we did not continue with the full search in PubMed. We did find some additional citations using this approach with PsychInfo, so we also repeated the citation search using this database. We compared the citations from our search to the citations included in the systematic review by Lucas-Thompson, Goldberg, and Prause (2010). As a result, we added the additional search terms "outcomes" OR "achievement" OR "adolescent" OR "adolescence," which increased our total to 1182 records. Duplicates were removed, reducing our total to 614 records.

Titles were screened for topic relevance by a research assistant and one author of this review. We excluded studies focusing on samples of children with disabilities from birth, premature children, or teen mothers. We excluded outcomes on usage of health services (i.e., vaccinations, nutrition related behaviors, doctor visits). Similarly, we excluded outcomes that are considered mediators leading to the final outcomes in our list (e.g., dietary intake, sedentary behaviors, sleep, work-family conflict). We excluded employment relating to migration or sex work. This screening provided us with 313 records.

For the final decision regarding inclusion, the abstracts and/or methods sections of papers were examined by an author of this review, with the other authors consulted regarding the empirical quality of the studies. We limited the studies to those published between 2002 and 2021. Studies were included if they had quantitative analysis, a measure of maternal work as a causal variable in the paper, at least one child or adolescent development outcome (cognition, schooling, behavior, anthropometrics, and health). Moreover, we restrict the analysis to studies analyzing natural or quasi-natural experiments and using the following methodologies: individual fixed or random effects, lagged dependent variables,[6] sibling fixed

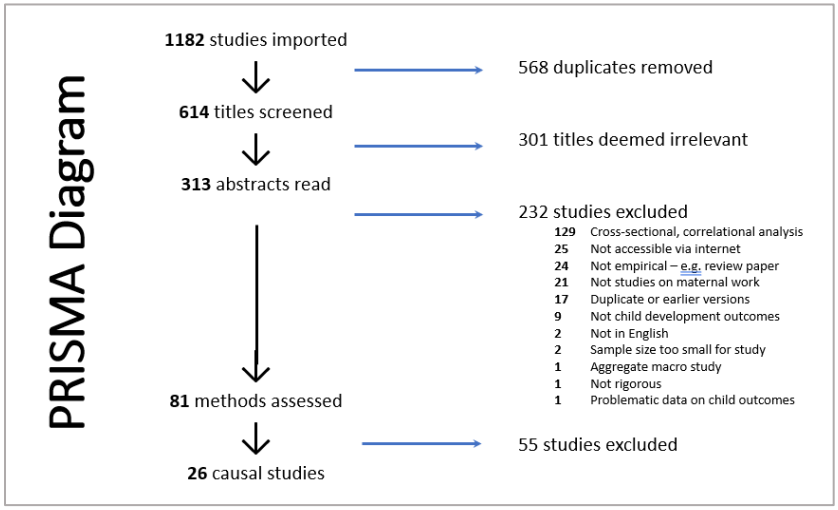
[5] The search was restricted to the titles of articles, since when the entire content of the articles was searched, it returned tens of thousands of papers. Several keyword searches (in contrast to title searches) of the permutations that yielded the most results did not provide additional papers for the review, so we did not pursue a keyword search.

[6] Even though many papers analyzed longitudinal data on maternal work, we excluded studies that only controlled for observables. One exception is the lagged dependent variable approach, which controls for the baseline value of the outcome variable. Although the lagged outcome variable is an observable, including it accounts for unobserved maternal or child characteristics that affected the earlier outcome, which reduces selection bias due to unmeasured child and family characteristics. However, if the study had controls for child development measures at baseline and these were from the same realm as the outcome variable, even if not exactly the same measure, we

effects, instrumental variables, difference in differences, or randomized control trials. We excluded therefore studies based on cross-sectional data (129),[7] non-analytical studies, such as review studies (24), macro studies (1), studies with small sample size (2), non-rigorous studies or with problematic data (2), studies not focusing on the selected child development outcomes (9), studies with the wrong intervention (21). Further studies were excluded because they were not written in English (2), not accessible via internet (25), or duplicates (17). This screening provided us with 81 studies.

We further limited our review to 26 studies using the most rigorous methods to establish causal impacts: those using instrumental variables, sibling fixed effects, and studies using multiple analytic approaches (most commonly individual fixed effects and lagged dependent variables). We did not find studies with difference in differences, randomized control trials, or studies that combined random effects with another method. Only one study relies on experimental variation generated by randomly assigned welfare-to-work program (Gennetian et al., 2005), however, it uses instrumental variables strategy due to imperfect compliance with program assignment. Three identification strategies used in the remaining causal studies include instrumental variables, fixed effects and lagged dependent variables (or some combination of these techniques).

Figure 1 Prisma Diagram



1.3 Scope of the analytical approaches used in the selected studies

Instrumental variables

The instrumental variables approach allows for identification of causal impacts of an endogenous independent variable (maternal work) on a dependent variable (children's development) when researchers can identify a variable which is strongly correlated with the former but not the latter (an instrument). The validity of instrumental variables relies on several assumptions, some of which are testable while others are not. First, there should be a strong correlation between the variables used as

did include the study in the review. For example, studies with the outcome BMI that controlled for birthweight (though not the dichotomous variable low birthweight) were included; studies on cognition that controlled for birthweight were not included if they did not also have an early measure of cognition.

[7] We excluded qualitative work and studies with cross-sectional or correlational analysis, including propensity score matching. Propensity score matching only uses observables to create comparisons and does not control for unobserved heterogeneity.

instrument and maternal work: as this assumption is testable, we only include papers with strong first stage in the review.[8]

The exclusion restriction – or assumption that the instrument only affects children's development through its impact on maternal employment only – is impossible to test. The range of papers that we use exploits a variety of instruments, from local unemployment rates to eligibility of the youngest child for kindergarten to changes in tariffs for female-dominated industries to child's health at birth. Exclusion restriction is more likely to be satisfied in some cases than in other. For example, children's birth weight is likely to affect children's development outcomes directly, in addition to maternal employment decisions. Several papers construct instrumental variables based on economic variables at the location of residence, such as local unemployment rates, female labor force participation, per capita governmental transfers. Variables reflecting local economic conditions may also affect children's development directly–for example, through higher quality educational establishments and medical services. Although we note that plausibility that exclusion restriction holds varies in papers we reviewed, we did not exclude any based on these criteria, as it is not testable, and exclusion would rely on our subjective judgement.

Fixed effects

Fixed effects strategies identify causal impacts through controlling for unobserved but fixed omitted variables. This approach relies on two core identifying assumptions. The first assumption is that conditional on a set of fixed time-invariant characteristics, maternal decision to work is as good as randomly assigned. The second assumption is that the effect of maternal work is additive and constant.

There are two variations of this identification strategy in the papers which we reviewed: a subset of papers uses maternal (or family) fixed effects, another set uses child fixed effects. In the former case, identification relies on comparing development outcomes of at least two children born to the same mother, when her work status varies by child. To identify causal impacts, we need to assume that unobservable differences between children do not affect maternal decision to work. In the latter case, researchers estimate impacts from comparison of outcomes of the same child overtime, with mother working in some periods and not working in other. Identification of causal impacts requires assuming that maternal decision to work does not respond to time-variant changes in child's development.

Assumptions in both cases can be easily violated: for example, mothers may choose to stay at home when their less academically able children are at school to help them with homework. Similarly, in the case of child fixed effects, mothers may adjust their work in response to children's development: for example, they may withdraw from the labor market in critical years, or in response to illness. Of course, the likelihood that the identifying assumptions are violated depends on the set of controls included in the estimating regression: for instance, it is more plausible that maternal decision to work in maternal fixed effects framework is as good as randomly assigned when conditional on her children's ability.

Lastly, fixed effects estimation is susceptible to attenuation bias from measurement error. Instrumenting independent endogenous variable can reduce the measurement error, and indeed several papers in our review rely on a combination of fixed effects and instrumental variables.

Several papers use random effects approach, which also relies on identifying causal impacts by purging the influence of time-invariant variables, but, unlike fixed effects, allows residuals for a given child to be correlated across periods.

[8] All selected papers either presented the first stage or stated that it was strong.

Lagged dependent variables

Lagged dependent variables also exploit variation in outcomes of the same child observed over time. However, identifying assumption is different: to establish causal relationship between maternal work and child's development in the lagged dependent variables framework, one needs to assume that maternal decision to work is as good as random, conditional on child's past development outcomes (not unobserved time-invariant characteristics). This assumption is likely to be violated if maternal decisions to work also respond to unobserved environmental variables, which may also affect children's outcomes such as employment of other household members.

Notably, fixed effects and lagged dependent variables can be thought of as bounding the causal effect of interest; and several studies included in this review include both FE and LDV. Consequently, we included studies that rely on lagged dependent variables when combined with child FE.

2. Samples

Our review covers 10 countries: 7 high income (Australia, Chile, Denmark, Italy, Norway, UK and US), and 3 lower-middle income (the Arab Republic of Egypt, India and Indonesia). Only four countries are non-Western: Chile, Egypt, India and Indonesia. Table A1 presents the distribution of the number of papers by country and method.

Table A1 <title>?>

Country	Number of studies	Number of papers by method			
		IV	Child FE & LDV	Mother FE	Mixed methods
Australia	1				1
Chile	1				1
Denmark	1	1			
Egypt, Arab Rep.	1	1			
India	2	1			1
Indonesia	1	1			
Italy	1	1			
Norway	1			1	
United Kingdom	3		1	1	1
United States	14	4	4	3	3
Total	26	9	5	5	7

The papers rely on a wide range of data, including survey and administrative data. A significant share of surveys are nationally representative, some focus on low-income populations. Some studies focus on a specific subsample – due to methodological demands (e.g., children with younger siblings only). Table AX <table 1?> presents the details for each study.

3. Scope of the measures

Maternal employment

There was a broad spectrum of how maternal employment was operationalized. 6 studies used exclusively a binary variable indicating whether the mother is employed and 3 studies differentiate between part-time and full-time employment with full-time usually defined as more than 35 hours per week.

Work intensity was considered in 13 studies and it was generally measured by continuous variables indicating weekly working hours[9] or employment duration, defined as the number of years, months, weeks or quarters in which the mother worked. In one instance (Reynolds et al. 2017) employment duration was measured as the number of weekly hours worked between the two surveys as a fraction of weekly working hours under full-time employment.

Many studies considered multiple employment variables simultaneously. For example, being employed and work intensity. Thus, the interpretation holds constant intensity while considering when the work occurred.

Finally, in 2 studies additional qualitative aspects of maternal employment were considered while controlling for working hours. Felfe and Hsin (2012) analyzed the impact of mothers' exposure to work-related hazards and stress on child cognitive and behavioral outcomes. Dunifon et al. (2013) considered standard in contrast to non-standard work schedules. These were measured as dummy variables, and could include night work, weekend work, or hours that changed weekly.

Child development outcomes

The studies reviewed cover a wide range of outcomes. To better organize the results, we grouped them in 6 categories: behavioral, cognitive, schooling, health, height and weight.

Behavioral outcomes are the most subjective as they are most often parent-reported often using a set of questions provided in the Child Behavior Checklist (CBCL), a validated checklist consisting of 99 items assessing a range of problem behaviors relating to socio-emotional well-being. In one study (Mendolia, 2016) different behavioral outcomes such as self-reporting on smoking, low self-esteem, high life satisfaction and intention to leave education were examined. Overall, 8 studies reported behavioral outcomes.

Cognitive outcomes, reported in 8 studies, vary as different tests for different realms (reading, mathematics, and receptive vocabulary, for example) are used. These were typically applied by a trained psychologist. The scores are generally not comparable across tests or children's ages (older children are expected to have higher scores than younger children), so results are standardized; sometimes standardization is done by child age and other times child age is a control variable.

Schooling outcomes were examined in 5 studies. The type of measures chosen differed according to whether children are observed during school age or not. In the first case, indicators of enrolment and being on track are used together with continuous variables such as the years of schooling (Dervisevic et al. 2022) or the time spent at school (Afridi et al., 2016). In the latter case, school achievement was measured by respondents' final grades at high school (Del Boca, 2016), achievement (yes/no) of advanced level qualification for admission in universities (Ermish and Francesconi, 2013) or by the years of education at age 27 (Haaland, Rege, Votruba, 2013).

Four studies included additional health outcomes. These cover a broad range from hospitalizations, injury and asthma episodes (Morrill, 2011) to parent reported perception of general health (Gennetian et al. 2010; Pekkurnaz, 2014) to more specific indicators such as levels of hemoglobin or lung capacity

[9] This was defined in most papers as average hours per week. Alternative definitions are present in two articles: Pekkurnaz, 2014 used the number of hours worked in the past week¸ Dunifon et al. 2013 recoded the number of hours worked into five <four?> dummies (1-19; 20-34; 35-44; 45+).

(Dervisevic et al. 2022). The breadth of sources from administrative data to surveys make this category of outcomes the most disparate.

Height outcomes include height for age z-scores (HAZ) and a binary variable for stunting, defined as low height-for-age, which is a proxy for chronic or recurrent malnutrition. Height is usually a direct observation, measured by a trained interviewer. Height measures cumulative nutrition and is particularly valuable prior to age 5, when growth distributions in healthy populations have been found to be similar worldwide (WHO, 2006). Three studies, all of them from non-Western lower middle-income countries use height outcomes (Rashad and Sharif, 2019; Shajan and Sumalatha, 2020; Dervisevic et al., 2022).

Weight outcomes are based on weight for age z-scores (WAZ) or body mass index (BMI), and – depending on the context – capture deviation from normal towards malnutrition or obesity.

Seven studies (Von Hinke Kessler Scholder 2008; Hubbard 2009; Anderson et al 2003; Bishop 2011; Pekkurnaz 2014; Greve 2011; Agiro and Huang 2020) considered binary indicators for obesity and overweight, which in most cases where constructed based on information on BMI. Two studies (Haaland, Rege, Votruba 201 and Ziol-Guest, Dunifon & Kalil 2013) use BMI. Three studies included other weight measures such as WAZ (Shajan and Sumalatha 2020) or an indicator of "underweight" (Dervisevic et al., 2022; Rashad and Sharif, 2019). Unlike height, weight fluctuates over time and therefore reflects current and acute as well as chronic malnutrition. The studies had a trained professional take this data.

4. Correcting for multiple hypotheses testing

Table 1: Bonferroni p-values and FDR q-values for papers with multiple outcomes and without multiple hypotheses corrections

Paper	Outcome	Coefficient	Standard error	Naïve p-value	Bonferroni p-value	FDR q-value
Morill (2011)	Overnight hospitalization	0.037	0.013	0.005	0.043	0.035
	Overnight hospitalization	0.045	0.017	0.009	0.081	0.035
	Overnight hospitalization	0.092	0.038	0.015	0.133	0.035
	Injury/poisoning	0.056	0.028	0.047	0.423	0.081
	Injury/poisoning	0.050	0.030	0.099	0.890	0.111
	Injury/poisoning	0.074	0.060	0.220	1.000	0.220
	Asthma episode	0.116	0.060	0.054	0.484	0.081
	Asthma episode	0.143	0.059	0.016	0.141	0.035
	Asthma episode	0.192	0.110	0.081	0.733	0.105
Rashad & Sharaf (2019)	Stunted	0.186	0.088	0.034	0.069	0.069
	Underweight	0.073	0.046	0.113	0.225	0.113
Shajan and Sumalatha (2020)	Height for age	-0.208	0.066	0.002	0.003	0.002
	Weight for age	-0.173	0.054	0.001	0.003	0.002
Felfe & Hsin (2012)	Letter word score	-0.055	0.024	0.022	0.066	0.059
	Passage completion	-0.049	0.026	0.059	0.178	0.059
	Applied problem solving	-0.048	0.025	0.055	0.165	0.059
Felfe & Hsin (2012)	Internal behavioral problems	0.048	0.025	0.055	0.110	0.110
	External behavioral problems	0.007	0.024	0.771	1.000	0.771
Von Hinke Kessler (2008)	Pre-school	0.004		Above 0.1	Above 0.3	Above 0.1

	Age 7	0.038		Above 0.01 and below 0.05.	Above 0.03 and below 0.15.	Above 0.03 and below 0.15.
	Age 11	0.019		Above 0.1	Above 0.3	Above 0.1
Furre Haaland (2013)	Years of education	0.013	0.005	0.012	0.050	0.012
	Completed high school	0.003	0.001	0.003	0.011	0.004
	Attended colleage	0.004	0.001	0.001	0.003	0.002
	Log inc 29	0.005	0.002	0.001	0.003	0.002
Afridi et al. (2016)	Time at school	6.506	1.102	0.000	0.000	0.000
	Enrollment	0.472	0.138	0.001	0.002	0.001
	Grade progression	0.406	0.101	0.000	0.000	0.000
James-Burmudy, 2005; weeks worked	PIAT reading, year 1	-0.00069	0.00064	0.281	1.000	0.422
	PIAT reading, year 2	0.00064	0.00059	0.278	1.000	0.422
	PIAT reading, year 3	0.00039	0.00047	0.407	1.000	0.488
	PIAT math, year 1	-0.00117	0.00067	0.081	0.485	0.422
	PIAT math, year 2	0.00034	0.00061	0.577	1.000	0.577
	PIAT math, year 3	0.0006	0.00051	0.239	1.000	0.422
James-Burmudy, 2005; hours worked	PPVT, year 1	0.03651	0.5222	0.944	1.000	0.944
	PPVT, year 2	-0.05182	0.05137	0.313	1.000	0.580
	PPVT, year 3	-0.00949	0.0469	0.840	1.000	0.944
	PIAT reading, year 1	-0.0526	0.02506	0.036	0.322	0.129
	PIAT reading, year 2	-0.02109	0.02131	0.322	1.000	0.580
	PIAT reading, year 3	0.01945	0.0864	0.822	1.000	0.944
	PIAT math, year 1	-0.07385	0.02595	0.004	0.040	0.036
	PIAT math, year 2	0.01512	0.02275	0.506	1.000	0.759
	PIAT math, year 3	0.04108	0.02026	0.043	0.383	0.129
Aughinbaugh and Gittelman (2003)	Any risky behavior	1.018	0.546	0.062	0.685	0.294
	Drank alcohol	0.629	0.246	0.011	0.116	0.116
	Smoked cigarettes	0.184	0.269	0.494	1.000	0.764

	Used marijuana	0.315	0.257	0.220	1.000	0.485
	Used other drugs	0.22	0.373	0.555	1.000	0.764
	Had sexual intercourse	0.058	0.243	0.811	1.000	0.892
	Convicted of crime	0.141	0.355	0.691	1.000	0.845
	drank alcohol at least several tiems a month	0.039	0.325	0.904	1.000	0.904
	smoked cigarettes everyday	0.438	0.315	0.164	1.000	0.452
	used marijuana at least 1-2 tiems per week	-0.511	0.489	0.296	1.000	0.543
	used no birth control at last sex	-0.544	0.311	0.080	0.883	0.294
Pilkauskas et al. (2018)	Externalizing behavior	-0.004	0.003	0.182	0.730	0.243
	Internalizing behavior	0.001	0.002	0.617	1.000	0.617
	PPVT	0.007	0.002	0.000	0.002	0.001
	Woodcock-Johnson	0.008	0.002	0.000	0.000	0.000

References

WHO (2006) WHO Multicentre Growth Reference Study Group: WHO Child Growth Standards: Length/Height-for-Age,Weight-for-Age,Weight-for-Length,Weight-for-Height and Body Mass Index-for-Age: Methods and Development. Geneva: World Health Organization.

Summaries of papers

9 IVS		Agiro and Huang 2020	Baum 2015	Del Boca et al 2016	Dervisevic, Lo Bue, & Perova 2021	Gennetian et al 2010
Publication		Journal of Family and Economic Issues	Journal of Labor Economics	European Journal of Population	The World Bank	Journal of Health Economics
Sample	**Data source**	The data for this study come from the Early Childhood Longitudinal Study, Kindergarten Class (ECLS-K)	NLSY	ISFOL-PLUS (Istituto per lo Sviluppo della Formazione Professionale dei Lavoratori Participation, Labour, & Unemployment Survey); child care center information from the Istituto degli Innocenti	Indonesia Family Life Survey (IFLS), SAKERNAS survey in 1995, and data on Indonesian tariffs from UN Comtrade.	National Evaluation of Welfare-to-Work Strategies Child Outcomes Study (NEWWS-COS)
	Years	2013	1986-1996	2008; 1992 data on the percentage of available places for public & private child care	1997, 2000, 2007/08, and 2014/15	1991-1994
	Sample size	1,591	4,944	12786	29,557 observations for the sample of children in two-parent households, and 32,214 observations in single mother households	3435
	Location	US	US	Italy	Indonesia	US (Grand Rapids, Riverside, Atlanta)
	Sample restrictions/ Mother characteristics	children with at least one younger sibling	Mothers who were between 23 and 30 in 1988. The low-income white oversample, the military sample, and children who did not live with their mother during their first three years were excluded from the analysis.		one sample is single mothers	welfare recipients
How employment is measured + additional characteristics of maternal work	**Employed or not**	employed or not	Whether the mother engaged in any marketplace work	Dummy variable for working or not. Could be full or part-time work.	employed or not	self-reports of employment
	Intensity	weekly work hours	Total number of hours worked in each week divided by 40 and the number of weeks in which the mother worked			number of quarters employed
	Other (eg. Standard schedule or not)		Number of weeks in which the mother worked full time and when the mother first started working after giving birth			
Children's Ages (in years unless otherwise indicated)	**Child ages when mother was working**	concurrent	1, 2 & 3 with Year 1 divided into quarters in some specifications	0-2	6-18	
	Ages at outcome measurement	in 2nd grade (ages 7-9)	tests given to children age 3 (PPVT) and age 5+ (PIAT-R & PIAT-M) (children an average age of 5 yrs)	19-30, however grades at the end of high school were when the student was approximate age 18	6-18	5-9
Methods	**Primary Causal**	IV (variation in youngest sibling's eligibility for kindergarten)	IV (local unemployment rate; potential earnings proxied by the local per capita income; the percentage of the local labor force that is female; the percentage of the local population that is urban and female; local population; percentage of the local population with a high school education and a college education; percentage of the local population employed; the percentage of the local labor force in manufacturing, in wholesale/retail trade, in services, in government work, and self-employed; and per capita government transfer payments)	IV (area of residence & regional supply of child care: number of child care centres, the slots available, and number of children admitted) These variables used as instruments for both child care attendance and maternal work.	IV (reduction in tarriffs in female-intensive industries)	IV (random assignment to an experimental welfare-to-work program)
Outcome: Schooling	**Measure**			high grades at the end of high school: dummy variable that equals 1 when the individual achieved the grade equal to or higher than 55/60 (for individuals who graduated before 1999), or 90/100 (for individuals who graduated after 1999, when a law changed the high school grading scale); otherwise it is zero. [Results robust to different grade cut-offs.]	years of schooling, currently enrolled, and on-grade ("age and schooling")	
	Findings			null	maternal work lowers the "age and schooling" indicator and increases years of schooling by around 1.8 years and the likelihood of being enrolled by around 35 percentage points with sufficiently low q-values to maintain significance with the FDR correction.	
Outcome: Cognition	**Measure**		PIAT-R & PIAT-M (5+ yrs) and PPVT (3-4 yrs) test scores. Test scores were averaged if there were multiple observations of the same test			
	Findings		when found, generally a negative impact of maternal work on cognition, but statistically insignificant in the IV models			
Outcome: Behavior/ socio-emotional	**Measure**					
	Findings					
Outcome: Height (haz)	**Measure**				height for age, stunting	
	Findings				children from working mothers, compared to childen from now working mothers, have better height for age score (around 1 standard deviation higher) and are around 49 percentage points less likely to be stunted	
Outcome: Weight (waz, bmi, obesity, overweight)	**Measure**	Overweight based on BMI (percentile used as robustness check)			wasting, underweight	
	Findings	null			null	
Outcome: Health	**Measure**				hemoglobin, lung capacity	Parent-reported child health scale (1-5) turned into a dichotemous variable for excellent or very good health in contrast to good, fair, and poor health
	Findings				children from working mothers, compared to childen from now working mothers, have higher (around 2.2 gr. per dl) levels of hemoglobin. Null results for lung capacity	Maternal employment reduces likelihood that child's health is reported as excellent of very good (as opposed to good, fair and poor)

IVS continued		Greve 2011	Morrill 2011	Rashad and Sharaf 2019	Shajan and Sumalatha 2020
Publication		Labour Economics	Journal of Health Economics	Oxford Development Studies	Journal of Public Affairs
Sample	Data source	Danish Longitudinal Survey of Children & administrative register data	National Health Interview Survey (NHIS)	Demographic and Health Survey (DHS)	National Family Health Survey (NFHS)
	Years	2003	1985-2004	2014	2015-2016
	Sample size	4336 mothers and 4348 children	274,842 children	12,888	37,557
	Location	Denmark	US	Egypt	India
	Sample restrictions/ Mother characteristics		children ages 7-17 with at least one younger sibling whose age range is around 5 years		
How employment is measured + additional characteristics of maternal work	Employed or not		maternal employment dummy variable for work in the past 1-2 weeks	current employment & employment in the previous 12 months	employed or not
	Intensity	average weekly working hours			
	Other (eg. Standard schedule or not)				type of employment considered in OLS specification
Children's Ages (in years unless otherwise indicated)	Child ages when mother was working	3 and 7	younger sibling is around age 5		concurrent
	Ages at outcome measurement		7 1/2 7-17, contemporaneous employment	0-5	0-5
Methods	Primary Causal	IV for maternal work at age 3 with local unemployment rate as instrument.	IV (youngest child's eligibility for kindergarten)	IV (cluster average of women's working status, with the exclusion of the women's own employment status)	IV (proportion of women in the district who work)
Outcome: Schooling	Measure				
	Findings				
Outcome: Cognition	Measure				
	Findings				
Outcome: Behavior/ socio-emotional	Measure				
	Findings				
Outcome: Height (haz)	Measure			Stunting	HAZ
	Findings			Maternal employment increases the probability of stunting by 18% .	IV estimates are -0.21 (HAZ) of a standard deviation lower for children of employed women.
Outcome: Weight (waz, bmi, obesity, overweight)	Measure	child overweight, defined by BMI cut-off above 97th percentile; robustness using the CDC cut-offs		Wasting. overweight & underweight	WAZ
	Findings	null		Maternal employment increases the probability of wasting by 13%. Null effects for overweight and underweight.	IV estimates are -0.17 (WAZ) of a standard deviation lower for children of employed women.
Outcome: Health	Measure		Hospitalized overnight (last 12 months), Asthma episode (last 12 months), Injury or poisoning episode (last 3 months)		
	Findings		Maternal employment results in an increase in the probabilty in overnight hospitalizations (4-9 percentage points), injury/poisoning (5-7 percentage points), asthma episode (11-19 percentage points)		

5 FE & LDV		Dunifon et al 2013	Felfe & Hsin 2012	Pekkurnaz 2014	Pilkauskas et al 2018	Von Hinke Kessler Scholder 2008
Publication		Developmental psychology	Economics of Education Review	Unpublished dissertation, The University o	Developmental Psychology	Health Economics
Sample	**Data source**	Fragile Families	PSID CDS, & O*NET	Early Childhood Longitudinal Study-Birth Cohort (ECLS-B)	Fragile Families	NCDS
	Years	2002 & 2004	1997 & 2002	2001/02, 2003/04, 2005/06, fall of 06, fall of 07	~2000, 2002, 2004, 2008	1958, 1965, 1969, 1974
	Sample size	2,367	1,630	1,450-8,900 depending on outcome	2011	3350
	Location	US (20 large cities)	US	US	US (20 large cities)	UK
	Sample restrictions/ Mother characteristics	Cohort recruited from hospitals, oversampling nonmarital births. Half of the sample is non-Hispanic Black, one quarter is Hispanic, and nearly one quarter is non-Hispanic White.	mothers who are working at the time of the survey		Cohort recruited from hospitals, oversampling nonmarital births.	excludes single parents
How employment is measured + additional characteristics of maternal work	**Employed or not**	Dummy variables for not working, working standard schedule, and working a non-standard schedule. The last group was divided into four non-mutually exclusive categories: weekends, nights, evenings, and different times each week.		full-time & part-time work dummies		Indicators for full-time and part-time for pre-school, age 7 and 11 employment
	Intensity	dummies for number of hours worked: 1-19; 20-34; 35-44; 45+	work hours	hours worked in the past week	number of months employed	
	Other (eg. Standard schedule or not)		work hazards & work stress		number of jobs working for 2 weeks or more	
Children's Ages (in years unless otherwise indicated)	**Child ages when mother was working**	concurrently for child FE, age 5 for LDV	past 5 years	concurrent	0-5	preschool, 7 & 11 for LDV or preschool, 7, 11 & 16 for FE
	Ages at outcome measurement	3 & 5	5-17	0-4	5 & 9	16 for LDV or preschool, 7, 11 & 16 for FE
Methods	**Primary Causal**	Child FE & LDV	LDV & Child FE	child FE & LDV (with LDVs instrumented with county- & state-level variables such as unemployment rate and poverty rate likely to affect working decisions, childcare prices and other factors influencing childcare availability)	Child FE & LDV	LDV (it controls for birth weight) & child FE
Outcome: Schooling	**Measure**					
	Findings					
Outcome: Cognition	**Measure**		Vocabulary, Reading, Applied problem solving	Cognitive ability (standardized composite i	PPVT & WJ	
	Findings		FE estimates significant and negative (-0.004- -0.005) for all outcomes considering an additonal hour worked. Slightly smaller & not significant in LDV specification. Hazards significant in all specificaitons, about -0.05 reduction in the standardized cognitive score.	FE & LDV both null	LDV: months employed between age 3 and 5 results in higher PPVT scores (0.005* SDs) and WJ scores (0.007** SDs) Consistent employment was not significantly different from fewer years of employment. FE: An increase in the number of months employed in the two-year period between ages 1 and 3, and ages 3 and 5, resulted in a 0.004* SD higher PPVT score and a 0.006* SD WJ score. Once the number of jobs were controlled, the coefficient on months of employment was no longer significant, although the point estimates were very similar.	
Outcome: Behavior/ socio-emotional	**Measure**	anxious/depressed and aggressive subscales of the CBCL	Internal & External Behavior Indices (Peterson & Zill)	Behavior problems (standardized composite index)	internalizing & externalizing subscales of the CBCL	
	Findings	Child FE: no significant impacts LDV: Mothers' night shift work is associated an increase in anxious behavior (0.21 sd) and aggressive behavior (0.13 sd) compared to not working; mothers with other types of work had children whose behavior had insignificant differences in comparison to children whose mothers were not working.	Results for behavior problems are null for hours worked. An increase in internalizing behavior (0.05) associated with work stress.	Behavior problems reduce -0.096 for each 10 hours worked in the LDV specification, FE null	Null results in all Child FE & LDV models	
Outcome: Height (haz)	**Measure**					
	Findings					
Outcome: Weight (waz, bmi, obesity, overweight)	**Measure**			Obesity & overweight		overweight (from BMI)
	Findings			Obesity reduces -0.035 for each 10 hours worked in the LDV specification, FE null		~5% increase in overweight at age 16 (LDV) and concurrently (FE) with full-time work at age 7
Outcome: Health	**Measure**			Health status (excellent or very good), ear infection, respiratory illness		
	Findings			Health status increases 0.029 with hours worked in the LDV specification, FE null		

5 Mother FE		Aughinbaugh and Gittleman 2004	Haaland, Rege, Votruba 2013	Waldfogel, Han, Brooks-Gunn 2002	Ziol-Guest, Dunifon & Kalil 2013	Ermish and Francesconi 2013
Publication		Journal of health economics	CESifo Working Paper	Demography	Social Science and Medicine	Journal of Applied Econometrics
Sample	Data source	NLSY79 & young adult supplement	Combined registry databases	National Longitudinal Survey of Youth (NLSY)	NLSY	British Household Panel Survey (BHPS)
	Years	1994, 1996, 1998, and 2000	1970-2007	1990, 1992, 1994, 1996	1994-2008	1991-1997
	Sample size	462 to 1414 depending on the outcome variable.	165,957 children in 77,581 families	548	2914 siblings from 1247 mothers	1026 (full sample); 647 (restricted sample)
	Location	US	Norway	US (nationally representative)	US	UK (nationally representative)
	Sample restrictions/ Mother characteristics	overly represented young mothers based on child age	Dropped families where siblings would have differential exposure to parental divorce or unusual living arrangements. Exclude children whose mothers or fathers died before the child reached age 16 and children who do not have siblings represented in the sample.	women in the NLSY who ranged in age from 16 to 32 at the time of their children's birth.		Full sample: children who: (i) are aged 18 or more and were born between 1970 and 1979; serious disabilities; (iii) lived with their biological, adoptive or step parent(s) during the first seven waves of the panel study; and (iv) have complete information employment patterns during childhood and other variables related to her. Restricted sample: full sample aged 16-17 when they lived with their parents
How employment is measured + additional characteristics of maternal work	Employed or not		Entry into full time employment	Employed during 1st year, 2nd or 3rd year, from age 3 up to the year before assessment, at time of assessment.		
	Intensity	Mother's hours of work were available for the first 3 years of the child's life and for the 3 calendar years preceding the adolescent interview		Hours worked during 1st year	average hours per week worked over child's lifetime	Number of months spent in full-time and part-time paid work. Worked at least 1 month when the child was aged 0-5
	Other (eg. Standard schedule or not)					
Children's Ages (in years unless otherwise indicated)	Child ages when mother was working	0-3 and the 3 years prior to the adolescent survey	0-18	1st year, 2-3 yrs, 4+ years	child's entire lifetime	0-5
	Ages at outcome measurement	15-18	27 or when enter military for boys	3-4 yrs: vocabulary; 5-6 yrs & 7-8 yrs: reading & math	13-14	aged 18 and more
Methods	Primary Causal	Mother FE	Mother FE	Mother FE	Mother FE	Mother FE
Outcome: Schooling	Measure		years of education at age 27, college attendance rates, log earnings age 29			dummy (yes/no) for having achieved an 'A (Advanced)-level' qualification or higher qualification
	Findings		5 additional years of full-time employment by one's mother reduces a child's education by 0.065 years, which amounts to 4 percent of a standard deviation in our sample. For high-school completion, college attendance & income, effects are significant but small, increasing by 1.5%, 2% and 2.5% respectively due to an increase of 5 years of mothers' work.			Negative and statistically significant effect of full and part time employment. Its point estimate ranges between 4 and 11 (full time) or between 3 and 7 (part time) percentage points lower probability for each additional year of full time employment, depending on the estimator.
Outcome: Cognition	Measure		IQ scores (boys only from military records)	Vocabulary (PPVT-R), Reading (PIAT) & Math (PIAT)		
	Findings		Insignificant impacts on IQ (boys only).	Primarily null; A significant (5% level) result is only found for the coefficient on reading (PIAT) ages 5-6 for employment during 2nd or 3rd year of life		
Outcome: Behavior/ socio-emotional	Measure	Incidence: Any risky behavior (of those that follow), drank alcohol, smoked cigarettes, used marijuana, used other drugs, sexual intercourse, convicted of crime. Thresholds: Drank alcohol several times a month or more, smoked cigarettes everyday, used marijuana at least once a week, no birth control at last intercourse				
	Findings	No impact of maternal employment during adolescence on participation in risky behavior. More hours of maternal employment during ages 0-3 is associated with an about 9 percentage point increase in the likelihood of ever having drunk alcohol with maternal FE model and a simlar likelihood of having used contraception the last time having sex with the grandparent FE model.				
Outcome: Height (haz)	Measure		height (boys only from military records)			
	Findings		impacts on height are insignificant (boys only)			
Outcome: Weight (waz, bmi, obesity, overweight)	Measure		BMI (boys only from military records)		BMI	
	Findings		impacts on BMI & are insignificant (boys only)		null	
Outcome: Health	Measure					
	Findings					

7 Mixed Methods		Afridi et al. 2016	Anderson et al 2003	Bishop 2011	Hubbard 2009 - Ch 1
Publication		IZA Journal of Labor & Development	Journal of Health Economics	Economic Record	Ph.D Thesis - The University of North Caro
Sample	Data source	YLS	NLSY	HILDA	ECLS-K
	Years	2007 & 2009-2010	biannually from 1986-1998	2,007	fall 1998, spring 1999, fall 1999, spring 2000, spring 2002, spring 2004
	Sample size	3,725	16,650 child-year observations from 6283 mothers; sample size for FE ranges from 4,159 for child FE to 7919 for sibling FE	907 (IV), 430 (mother FE)	18,990 person-years for 6,330 indivduals
	Location	India (state of Andhra Pradesh)	US	Australia	US
	Sample restrictions/ Mother characteristics	Sample restricted to households in rural areas in both periods			
How employment is measured + additional characteristics of maternal work	Sample restrictions/ Mother characteristics	Dummy taking value 1 if the mother participates in the labor market and 0 otherwise			
	Employed or not		average hours per week (when working) and total weeks worked	number of years working full and part-time	mother's employment hours, divided into not working (0-8), part-time (9-35), full time (36+)
	Intensity				also if in childcare>5 hours per week
Children's Ages (in years unless otherwise indicated)	Other (eg. Standard schedule or not)	concurrently	cumulative from beginning of life (although one simple probit considers work before & after age 3)	12-15 (The exception is 19-year-olds, for whom adolescence is the 3-year period between the ages of 11, 13 and 15, as necessitated by the 7-year panel)	concurrently
	Child ages when mother was working	5--14	between 3 & 11; varies based on specification	15-19	kindergarten, 1st, 3rd & 5th grades (approx ages 5, 8, & 10)
Methods	Ages at outcome measurement	Child FE & IV. Instruments used: lagged rainfall and lagged NGRES funs sanctioned at the beginning of each financial year. Both instruments are at the mandal (subdistrict) level.	Child FE, Mother FE, IV (unemployment rate, child care regulations, wages of child care workers, welfare benefit levels, and the status of welfare reform)	IV (local unemployment rate by industry and typical hours worked by women in that industry locally) & mother FE	FE; some specifications have lagged outcome variable; Discrete Factor Random Effects (with IV). Instruments that are state or county & time variant: 7 employment-related variables & 6 child-care-related variables; also instrument with child's birth weight
Outcome: Schooling	Primary Causal	time spent at school; school enrolment; grade's progression (defined as the actual grade attainment of a child divided by the grade the child should have completed at her/his age)			
	Measure	The authors find that if the mother works, her child's time spent in school (significantly) increases by 6.506 h in a day. The probability of a child being enrolled in school (significantly) rises by 47.2 percentage points. Moreover, there is a significant effect of mother's work participation on her child's grade progression: the gap between a child's actual and ideal grade declines by 40.6 percentage points.			
Outcome: Cognition	Findings				
	Measure				
Outcome: Behavior/ socio-emotional	Findings				
	Measure				
Outcome: Height (haz)	Findings				
	Measure				
Outcome: Weight (waz, bmi, obesity, overweight)	Findings		overweight, determined by the CDC's BMI guidelines	BMI & Overweight	overweight & obese, as determined by BMI
	Measure		Child FE: children of mothers who work an additional 10 hours per week while working face a 1.5 percentage point increase in the likelihood of being overweight. IV: Null	Not statistically significant relationship with both full & part time employmentin the IV & sibling FE specifications	When mothers work full-time and do not use child care, the risk of obesity decreases by 1.5 percentage points and decreases risk of being overweight by around 2 percentage points
Outcome: Health	Findings				
	Measure				

Mixed Methods, continued		James-Burdumy 2005	Mendolia 2016	Reynolds, Fernald & Behrman 2017
Publication		Journal of Labour Economics	Journal of Family and Economic Issues	Social Science and Medicine - Population H
Sample		NLSY	British Youth Panel (BYP), a complement to the British Household Panel Survey (BHPS)	Longitudinal Survey of Early Childhood (ELPI)
	Data source	1986 & 1988	1994-2006, yearly waves	2010 & 2012
	Years	498 (PPVT), 1,761 (PIAT)	around 5,000 (number of observations reported, not number of children)	2476
	Sample size	US	UK	Chile
	Location	mothers ages 21-27 in 1986	employed mothers	Working mothers
How employment is measured + additional characteristics of maternal work	Sample restrictions/ Mother characteristics			
	Employed or not	hours worked in the 1st, 2nd & 3rd year of life	hours per week	Fraction worked of the two-year period between surveys, calculated from both hours work and employment duration
	Intensity			
Children's Ages (in years unless otherwise indicated)	Other (eg. Standard schedule or not)	0-3	concurrent	1-3 years
	Child ages when mother was working	5-18 for PIAT, 3-5 for PPVT	11-15	1 & 3 years
Methods	Ages at outcome measurement	Mother FE and IV (percent of county labor force employed)	Child FE & Mother FE	LDV, IV (mothers employed before child born, percent of other mothers with children the same age working part & full time, community gender equality index and traditional values index)
Outcome: Schooling	Primary Causal			
	Measure			
Outcome: Cognition	Findings	PPVT (ages 3-5) & PIAT reading & math (ages 5-18)		PPVT, Battelle
	Measure	Average hours worked per year in first 3 years of child's life do not significantly affects child's PPVT score. Analogously, no significant effect of hours worked in the first, second or third yeard of the child's life on PPVT scores and PIAT readings scores. Only hours worked in year 1 are associated with lower PIAT math scores: Working 1 additional 8-hour day each week in year 1 (an increase of 416 hours for the year) would reduce the PIAT math score by 0.5 points relative to the mean of 96.		null
Outcome: Behavior/ socio-emotional	Findings		youth reports of: smoker, low self esteem, high life satisfaction, and intention to leave education at age 16	Behavior (CBCL)
	Measure		null	null
Outcome: Height (haz)	Findings			
	Measure			
Outcome: Weight (waz, bmi, obesity, overweight)	Findings			
	Measure			
Outcome: Health	Findings			
	Measure			