Still Flying: Reply to "Not a Flying Start after All?" by Lillebø et al.

Pedro Carneiro

University College London, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, and Norwegian School of Economics–Centre for Experimental Research on Fairness, Inequality, and Rationality

Katrine Løken

Norwegian School of Economics–Centre for Experimental Research on Fairness, Inequality, and Rationality

Kjell G. Salvanes

Norwegian School of Economics–Centre for Experimental Research on Fairness, Inequality, and Rationality

I. Introduction

In Carneiro, Løken, and Salvanes (2015), we analyze the long-term impacts of the 1977 Norwegian maternity leave reform on several outcomes of children. Lillebø et al.'s (2024) comment presents several objections to our analysis. The most important part of their comment uncovers details of the reform that were missed in our paper. In addition, they also document a coding error and raise a few objections to the empirical work we conducted.

Lillebø et al. have conducted excellent archival work, and their corrections to our paper's description of the details and implementation of

We thank the referees and the editor, Lance Lochner, for helpful comments on preparing the reply.

Electronically published October 25, 2024

Journal of Political Economy, volume 132, number 12, December 2024.
© 2024 The University of Chicago. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits non-commercial reuse of the work with attribution. For commercial use, contact journalpermissions@press.uchicago.edu. Published by The University of Chicago Press.
https://doi.org/10.1086/732220

the reform are very important. We deeply regret having missed the issues raised by Lillebø et al. when we worked on the paper, and we are happy that this can be corrected. We now present new estimates of the impact of the reform that address all our errors.

In the remainder of the reply, we discuss the following points:

1) The facts uncovered in the comment could have implications for the research design used in our paper. They argue that it is impossible to use such a research design, and we agree that, in theory, they have a valid point. This led us to revisit all the details of our original work and question all our results. Whether Lillebø et al.'s arguments invalidate the research design in our paper depends on the extent to which individuals understand all the complex details of the reform (as opposed to having a more simplified view of its most salient details) and how they react to them. Below we explain under what conditions the research strategy in our paper could, in theory, still be valid, and we present evidence that this is likely to be the case. In particular, we discuss sensitivity checks originally shown in our paper as well as some new ones developed for this reply, which strongly suggest that the original results were robust to the information uncovered in the comment. Perhaps more importantly, even if the authors were correct that our paper's research strategy is completely invalid, it is possible to redesign a research strategy, modified from our original paper, that is immune to the authors' criticisms. We present new estimates of the impact of the reform using this new research strategy, showing estimates that are similar to the original results in our paper. These can be taken as the best estimates of the impacts of this reform on the long-term outcomes of children.
2) A second implication of the comment for the work in our paper, which is important and also true for our new results, concerns the interpretation of the estimated impacts of the reform. The authors document that, compared with the prereform regime, the new maternity leave entitlements were less generous than we claimed them to be, and they are right. That said, the reform still provided a substantial increase in maternity leave benefits for the majority of mothers. We fully incorporate their point in the interpretation of our new estimates.
3) In addition, the comment raises objections to some aspects of the treatment of the data in our paper, including a minor coding error. None of these issues have any implications for our estimates, as we show; we briefly discuss one issue here and relegate answers to these objections to the appendix, available online.

To simplify the reading of the comment and this response, the remainder of our reply follows the structure of Lillebø et al. (2024) section by section.

II. The Contents and Implementation of the Maternity Leave Reform—Implications for Our Paper

Section II of the authors' comment presents a more complete description of the reform than the one provided in our paper. Their comment conducted excellent research on legal and newspaper archives, which was enabled by their access to newer archival search tools. For instance, nowadays newspapers and other documents are digitized, making it easier to search, an option that was not available to us years before (all our archival searches were manual). It is difficult to get all the details correct, including changes due to labor law and social security law, as well as those arising from biannual union negotiations varying by the industry level.

Importantly, not even Lillebø et al. (in their more recent and more thorough analysis) can have full certainty about all the details of the reform and its implementation due to limited data availability and documentation for this period. For example, none of the official documentation has information about the exact eligibility criteria.[1] To the best of our knowledge, our research was thorough and covered the most important aspects of the reform, and we also engaged with several other social scientists in Norway. In hindsight, it is clear where we should have dug deeper in our archival research.

Among the new details of the reform described by the authors, there are two that are especially important and potentially have implications for the results and interpretations in our paper. These are also the ones emphasized in the comment. First, the reform was less generous than described in our paper. Second, the comment provides new information on different dates of implementation for different groups of workers and a transitional arrangement overlooked in our paper, both of which potentially allowed some women access to the new maternity leave benefits even if they delivered their children before the date of the implementation of the reform, July 1, 1977.

Regarding the first point, we originally had some uncertainty about the exact maternity leave entitlements available to mothers before the reform (reflected in slight changes in the description of the reform in different working-paper versions of our paper; Carneiro, Løken, and Salvanes 2008, 2010). At the time, we did not have available the extensive material that

[1] The best-verified information about the exact eligibility threshold is found in a newspaper article after the digital search (see fig. 8; all figures are available online).

the authors had uncovered and we were convinced that, by and large, the reform represented a change from 12 weeks of unpaid leave to 18 weeks of fully paid leave. In addition to the actual 1977 law, the main government documents that described the reform, from trusted sources such as Statistics Norway (1991) and Norges Offentlige Utredninger (NOU 1996, 27), did not include many relevant details and referred to the July 1 reform.[2] Lillebø et al. revealed instead that the reform consisted of a change from 12 weeks of partially paid leave to 18 weeks of partially paid leave. We have provided additional documentation (see fig. 2) that this partially paid leave was in practice close to full coverage for most workers.[3] However, there was no change in the amount of monthly maternity leave payments on July 1, 1977; the only change was the length of allowed leave. The fact that the reform was less generous than we described requires us to reinterpret our results as an extension of 6 weeks of paid leave for mothers (not an extension from 12 weeks of unpaid leave to 18 weeks of paid leave). That said, we should note that this constitutes a substantial increase (of 50%) in the amount of paid leave entitlements.

Regarding the second point, we were unaware that the reform was implemented on different dates for the private and public sectors. We described the reform as affecting all births after July 1, 1977, whereas in reality this was true only for workers in the private sector (workers in the public sector benefited from an earlier arrangement from union negotiations for public-sector workers as documented in the comment). Since private-sector workers were the majority in the labor force (in fig. 3, we document that around 70% of women worked in the private sector in the 1970s), this oversight, even if unfortunate, may not have had substantial practical implications for our research. If anything, this will dilute the impacts of the reform by including some ineligible mothers in the sample.

More serious was our failure to appreciate the importance of the transitional agreement, which could potentially give access to the new benefits to women delivering up to 12 weeks before the July 1 date. Almost all the information we had access to, from both public-sector reports and the media, emphasized the July 1 date as the important date of the reform.[4] We were aware of the possibility that there could be a transition arrangement in place, but in the documents we had access to it was unclear whether

[2] For example, fig. 1 from NOU (1996, 27) refers to July 1, 1977, as the main reform with 18 weeks of 100% income replacement. It also refers to the 12 previous weeks as having some payment but no information on whether it was universal and what the amount was.

[3] About 75% of the labor force had 90%–100% coverage through different agreements between the state and labor unions. The remaining 25% was covered by the public system (Folketrygden), and compensation varied from 40%–90% based on income and length of the spell.

[4] Figure 4 shows the newspaper advertisement inserted by the government's information service on June 30, 1977. Here the date July 1, 1977, is mentioned seven times, including in the first paragraph.

there was one, and it was never mentioned as being very important. We also believed that even if such a transitional arrangement existed in theory, in practice it did not affect our research design because it was not practical to take advantage of it and because the salient date of the reform was July 1.[5] We had included versions of this argument in an early working paper (Carneiro, Løken, and Salvares 2008), but unfortunately this reference got lost during the many iterations of the publication process. Arguments that support the assertion that July 1 was very salient and may have been in practice the date to which most mothers reacted are summarized as follows: (1) the July 1 date was especially salient in all the media reports of the reform, (2) the public was not made fully aware of the transition arrangement until close to the reform date, and (3) it is likely that women giving birth after April 9 but before July 1 had already exhausted their maternity leave (under the previous regime) many weeks before the law came into place, were back to work, were surprised by the details of the transition arrangement, and could not easily take advantage of it.[6]

The absence of take-up data for this reform prevents Lillebø et al. or us from showing whether mothers took up new benefits differentially if their delivery date was just before or just after the July 1 date. In theory, even with the arguments presented above about salience, the transition arrangement is a threat to the research strategy in our paper, as rightly emphasized by the authors. However, we show that it is possible to modify the original paper's research design in a way that fully accounts for the presence of the transition agreement and that we use to produce new estimates of the impact of the reform.

Our new estimates, which are robust to the existence of the transition arrangement, are similar to those in the original paper. This suggests that our assumptions were not far off the mark. Furthermore, our original paper documents (in fig. 8*B*) that our results were robust to substantial changes to the research design, including the omission of several weeks of data on either side of the July 1 date (we did it to have estimates robust to potential nonrandom manipulation of dates of birth), which lends further support for the idea that the original research design was valid after all.[7]

[5] It is not clear to us when the transition arrangement became public information. It was likely announced at the same time that the law was approved, on June 10, 1977, although it probably was not given nearly as much prominence as the law itself, and we could not find any mention of it in newspapers until June 30, 1977. In fig. 5, we show an example that even after the reform date, the government was actively seeking new mothers to make them aware of the full range of new benefits now available to them, suggesting that many of them (namely, those giving birth before July 1) may not have been fully informed of all their benefits.

[6] They also had to inform their employers of plans (see fig. 6), which could have complicated uptake of the transitional benefit.

[7] Since the publication of our paper, two additional papers have used similar research strategies to examine the impacts of this maternity leave reform on different outcomes: Bütikofer, Riise, and Skira (2021) and Schwartz (2021). Results from both of these papers suggest that the July 1, 1977, date was indeed important.

The new estimator we suggest is a modification of the regression discontinuity–difference-in-differences (RD-DD) estimator in the comment, which begins by comparing children born on July 1, 1977, or later (treatment) and children born on April 8, 1977, or earlier (control). This is the first difference of the DD estimator (it is not quite an RD estimator anymore). We compute this same difference for births in a comparison year (e.g., 1975), and then we take the difference in these two differences. This is a straightforward adaptation of our original (RD-DD) strategy, excluding from the sample all births occurring between April 9 and June 30, both in 1977 and in the comparison years. It incorporates the new information in the comment and takes a conservative view of the implementation of the transition arrangement by excluding from the analysis all births occurring in the transition period. In our view, it is the best empirical strategy to analyze this reform. See appendix B for a formal exposition of the model, in addition to an alternative specification that incorporates data from the transition period (fig. 10; online table 1, panel B).

Table 1 shows the estimated impacts of access to expanded maternity leave benefits on dropout rates, college, and the log of earnings at age 30, using the original version of the data used in our paper (see also fig. 9). In

TABLE 1
ALTERNATIVE SPECIFICATION TO ACCOUNT FOR TRANSITION AGREEMENT, APRIL 9–JUNE 30, 1977

Variable	Dropout Rate	College Attendance	ln Earnings Age 30
	A. Baseline—A Flying Start		
Estimates:			
RD-DD years	−.019**	.020*	.050***
Controls: 1975, 1978, 1979	(.007)	(.011)	(.016)
	[.19]	[.44]	[12.5]
Observations	63,571	63,571	60,732
	B. Drop April 9–June 30		
Estimates:			
RD-DD years	−.017**	.013	.037**
Controls: 1975, 1978, 1979	(.008)	(.011)	(.016)
	[.19]	[.44]	[12.5]
Observations	61,464	61,464	58,685

NOTE.—Panel A shows the baseline estimates from the original paper. Panel B shows the modified estimator excluding the transfer period between April 9 and July 1. Standard errors are shown in parentheses, and the mean of the dependent variable for the months before the reform is shown in brackets. Finally, we include the total number of observations. We follow eq. (1) with a window of 90 days to the left of April 9 and to the right of July 1. We include a linear trend and triangular weights. We allow the trends to vary on each side of the discontinuity; however, we impose the same trend across years. The treatment year is 1977, and the control years are 1975, 1978, and 1978.

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

panel A, we replicate the findings in our original paper (which are identical to those in that paper's table 3). In panel B, we implement our new and preferred research design on the same data and compare children born on July 1 or after with children born before April 9. Figures 9 and 10 also show parallel trends in outcomes between those born in the pretreatment months during the treatment year and those born during the control years.

The impacts on high school dropout are very similar across panels A and B. The impact on college attendance is slightly lower in panel B,[8] and the impact on log earnings at age 30 is 3.7% in the most robust models, as opposed to 5% in the original paper.

In sum, the most important issue raised by Lillebø et al. for the validity of the research design in our paper is the fact that the 1977 maternity leave reform in Norway allowed for a transition period. This meant that women who were deemed ineligible for the reform in our paper could potentially access the new benefit, although this depends on how the reform was perceived by mothers. Here we present a simple modification to the research design in our paper, which fully addresses this issue and omits from the data individuals born during the transition period. The results we generate from our corrected procedures are very similar to those in the original paper, suggesting that the original research strategy was valid after all.

III. Background and Press Coverage—Do They Affect the Timing of Births?

The additional press coverage that Lillebø et al. provide shows that this reform was important beyond what we documented in the original paper. But it took many years from the time it was prominent in newspapers to actual implementation, and there is much less discussion in the newspapers about the details of implementation, such as the relevant dates, making it very hard for mothers to time their births and the take-up of these benefits in response to the exact reform date (which was the main argument in our paper when we argued that reform was not prominent in newspapers early on). See figure 7 for a longer reply to this subsection.

IV. Some Minor Issues with Modeling Choices and Empirical Implementation

In the last section of their comment, Lillebø et al. present several disagreements with the statistical analysis in our paper. These disagreements are

[8] In the original paper, we showed that effects on college attendance were not very robust. This is also seen in fig. 9 showing the raw data.

mostly independent of the new description of the reform and its implementation. They concern sample restrictions, choice of control years, and the potential use of different variables. They also point out a coding error.

The comment's first set of claims is about our choice of control years, and we briefly discuss this here. They argue that 1976 should not have been excluded from the control years. We had argued that in 1976 there was an abortion reform that potentially hit births around the July 1 cutoff that made 1976 a poor control. They claim that the abortion reform was not relevant and that 1976 is a valid control year. In addition, they also suggest that 1978 should not be included as a control because there was a change in maternity leave payments for mothers delivering their children before and after July 1, 1978.

These are reasonable arguments, and we took them seriously. We show that 1976 is not a valid control year since, when comparing the characteristics of mothers giving birth before and after July 1 in different years, we found an imbalance in the characteristics of mothers in 1976 but not in any other years (see fig. 14). We suggested that this could be due to the abortion reform, but we cannot be sure of that. Regardless of the source of the imbalance, it is there. Concerning the 1978 change in maternity leave benefits, we argue that it is too small in practice to make a difference.

Ultimately, these concerns do not have any substantial effect on the estimates in our paper. This can be seen in tables 2–5 (tables 2–12 are available online), which present our estimates with and without the inclusion of 1976 as a control year and with and without the inclusion of 1978 as a control year. The results are similar to the baseline estimates excluding 1976 and including 1978.

In addition, it is true that, unfortunately, we made a coding error in defining eligibility. That said, as shown in tables 6 and 7, it has little consequence to our results. This is also the case for a couple of more minor objections raised in the comment.[9]

V. Conclusion

Lillebø et al. (2024) reexamines the evaluation of the long-term impacts of the 1977 maternity leave reform conducted in our paper. Their archival work showed that we had made mistakes in the description of the reform and its implementation. These potentially affected the research design used to evaluate the reform as well as the interpretation of the estimated impacts.

[9] Figures 11 and 12 show old results from the original paper and from a previous working-paper version (Carneiro, Løken, and Salvares 2010), further suggesting that our results are quite robust. Figure 13, also from the original paper, shows that the eligibility definition we used matches well with official labor force statistics. Tables 8–12 show the robustness of our findings to using different measures of earnings.

The new facts uncovered by the authors are important and led us to reestimate the impacts of the reform using a new research strategy that takes these facts into account. Accordingly, we present new estimates of the impacts of the reform, as well as a more correct interpretation of their meaning.

Our improved estimates show that the maternity leave reform in Norway had large long-term impacts on the lives of children. Quantitatively, they turn out to be similar to the original estimates in our paper. We conjecture that this is because the original research strategy may have been adequate, and we explain why. However, this did not have to be the case. The estimated impacts under the corrected empirical strategy could have turned out to be very different from the original paper's results.

In addition, we need to make one important correction to the interpretation of the results that we got wrong. The reform did not lead to an expansion in maternity leave benefits from 12 weeks of unpaid leave to 18 weeks of paid leave. The actual reform was quite significant but not as dramatic and consisted of an expansion of maternity leave benefits from 12 weeks of paid leave to 18 weeks of paid leave. This means that extension of leave, at least from a low base, can have long-term impacts on children. There are several reasons why there could be positive effects from this reform compared with what other papers on extensions have found. The counterfactual care (e.g., maternal care, formal care, or informal care), the weeks of prior leave before the reform, the labor market conditions at the time of the reform, and the health care system supporting mothers and infants are all candidates that could vary between countries and over time. The best evidence we have for a mechanism is the findings in Bütikofer, Riise, and Skira (2021) that the reform affected the health of the mother (the other papers in this literature have not been able to study the effects on maternal health). A healthier mother could be key for the children also in the longer run beyond the extra weeks spent with the mother in early life.

References

Bütikofer, Aline, Julie Riise, and Meghan M. Skira. 2021. "The Impact of Paid Maternity Leave on Maternal Health." *American Econ. J.: Econ. Policy* 13 (1): 67–105.

Carneiro, Pedro, Katrine V. Løken and Kjell G. Salvanes. 2008. "A Flying Start? Maternity Leave and Long-Term Outcomes for Mother and Child." Working paper, Dept. Econ., Norwegian School Econ., Bergen, Norway.

———. 2010. "A Flying Start? Long Term Consequences of Maternal Time Investments in Children during Their First Year of Life." Discussion Paper no. 5362, Inst. Labor Econ., Bonn, Germany.

———. 2015. "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children." *J.P.E.* 123 (2): 365–412.

Lillebø, O., S. Markussen, K. Roed, and Y. Zhao. 2024. "Not a Flying Start after All?" *J.P.E.* 132 (12): XXX–XXX.

NOU (Norges Offentlige Utredninger). 1996. "Offentlige Overføringer til Barnefamilier." Report no. 1996:13, Norges Offentlige Utredninger, Oslo. https://www.regjeringen.no/contentassets/5a5c94a53b474bc3bedf00d9f836eee4/no/pdfa/nou199619960013000dddpdfa.pdf.

Schwartz, Mariel. 2021. "The Spillover Effects of Maternity Leave Policy on Young Women's Schooling Choices." Working paper, Dept. Econ., Univ. Chicago.

Statistics Norway. 1991. "Oversikt over Endringer fra Begynnelsen av 1960 Tallet i Lover og Regelverk for Utvalgte Emner i Tilknytning til Barn." Internal Note no. 91/17, Statis. Norway, Oslo.