Supplemental Appendix

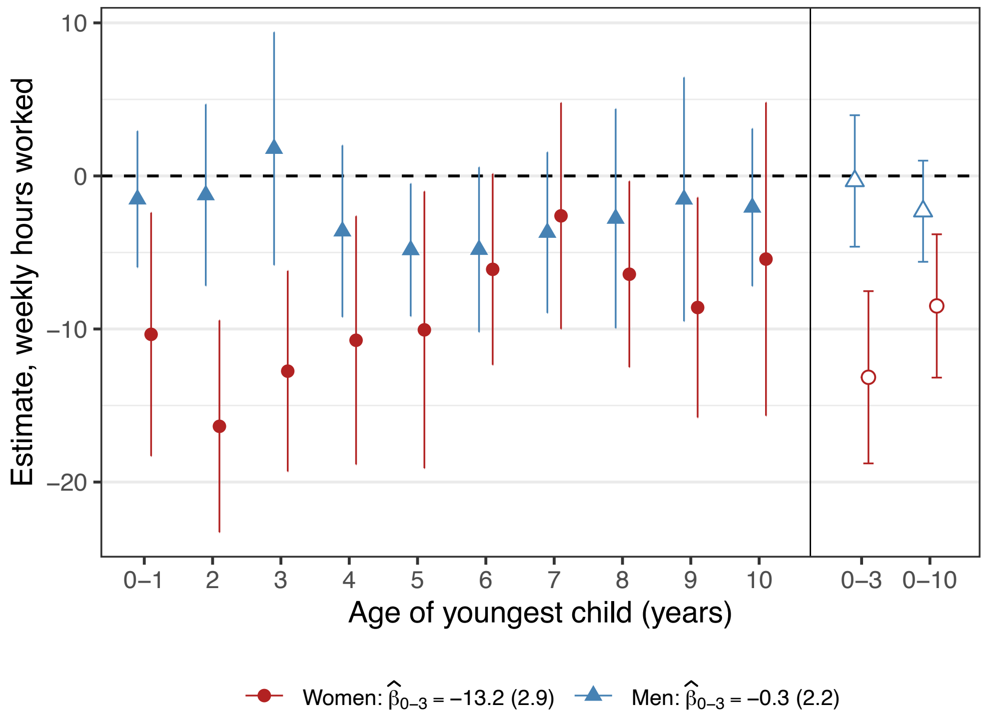
A Additional Exhibits

Table A1: Estimates for female owners under different weighting/outcomes

	(1)	(2)	(3)	(4)
	Response as percentage change			
$t = -1$	–0.010 (0.027)	–0.014 (0.025)	0.004 (0.025)	–0.022 (0.044)
$t = 0$	–0.073 (0.032)**	–0.075 (0.029)**	–0.066 (0.030)**	0.034 (0.057)
$t > 0$ avg.	–0.230 (0.043)***	–0.207 (0.038)***	–0.205 (0.038)***	–0.215 (0.063)***
	Response in levels			
$t = -1$	–820 (2,279)	–1,120 (1,981)	294 (1,963)	–943 (1,868)
$t = 0$	–6,251 (2,869)**	–6,178 (2,501)**	–5,388 (2,499)**	1,467 (2,434)
$t > 0$ avg.	–22,088 (5,247)***	–18,588 (4,198)***	–18,283 (4,194)***	–11,219 (4,090)***
Outcome variable	Profits before owner salaries	Profits before owner salaries	Profits before owner salaries	Operating profits
Weighting vars.	Firm age × SIC section	Firm age × SIC division (2-digit industry)	Firm age × SIC section × owner educ.	Firm age × SIC section
$\bar{Y}$	82,581	78,485	78,511	40,467
N owners	45,902	42,236	39,390	45,901
N firms	54,856	49,597	46,332	54,856

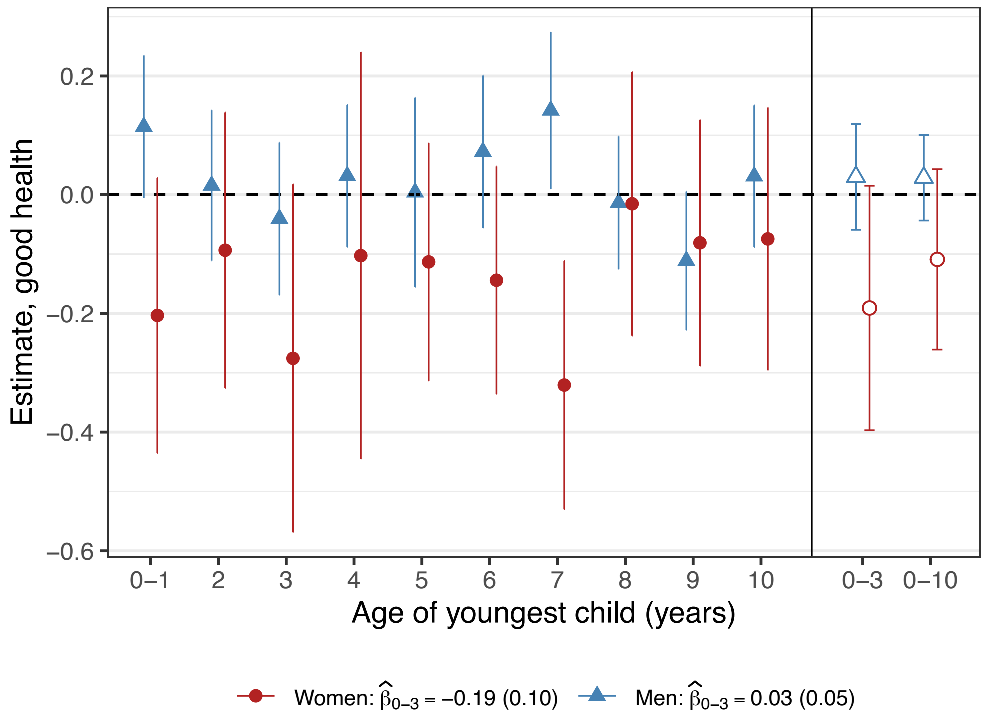
Notes: This table reports estimates of the response of profits to firm owner child birth for female-owned firms under different estimation weights and outcome variables. Column (1) reproduces our estimates from Figure 1 and Table 2. Columns (1)-(3) use our preferred profit outcome, which is total revenues minus all costs except for salaries paid to firm owners. In column (4), we replace this outcome with total revenues minus total costs, where costs now include the salary paid to the owners (i.e., the outcome is operating profits). In columns (1) and (4), we follow our approach from the text and weight each control cohort to match the joint distribution of firm age and firm industry (SIC section) in the corresponding treated cohort. In columns (2) and (3), we replace the SIC section with more granular measures of industry. Column (2) weights based on SIC division, which is comprised of 2-digit industries. Column (3) weights based on the interaction between SIC section and owner education (3 levels). We drop observations with values of weighting variables for which there is no overlap between treatment and control, which reduces our sample size in columns (2) and (3). All other notes are as in Figure 1 and Table 2.

Figure A1: Entrepreneur working hours by age of youngest child



Source: European Union Statistics on Income and Living Conditions (EU-SILC), Norway, 2008-2023.
Notes: This figure plots regression estimates of weekly hours worked on the age of the entrepreneur's youngest child, along with 95% confidence intervals based on robust standard errors. Regressions are estimated separately by the sex of the entrepreneur. The control group is comprised of entrepreneurs with no children. The regressions include dummies for calendar year, firm industry, the marital status and education level of the entrepreneur, and a quadratic in entrepreneur age. The "0-3" estimate, which is also reported in the legend, is an average of the age 0-1, 2, and 3 estimates. The "0-10" estimate is an average over all ages. The sample includes self-employed individuals who have at least one employee (N = 224 women and N = 647 men).

Figure A2: Entrepreneur health by age of youngest child



Source: European Union Statistics on Income and Living Conditions (EU-SILC), Norway, 2008-2023.
Notes: This figure plots regression estimates of a "good health" indicator on the age of the entrepreneur's youngest child, along with 95% confidence intervals based on robust standard errors. The outcome takes a value of one if the individual reports being in excellent or very good health, and zero otherwise. Regressions are estimated separately by the sex of the entrepreneur. The control group is comprised of entrepreneurs with no children. The regressions include dummies for calendar year, firm industry, the marital status and education level of the entrepreneur, and a quadratic in entrepreneur age. The "0-3" estimate, which is also reported in the legend, is an average of the age 0-1, 2, and 3 estimates. The "0-10" estimate is an average over all ages. The sample includes self-employed individuals who have at least one employee. The sample includes self-employed individuals who have at least one employee ($N = 224$ women and $N = 647$ men).

B Estimation

This section provides a more precise description of our estimation procedure. We perform this procedure completely separately for men and women. This conditioning remains implicit throughout the remainder of this section.

Let $\mathcal{S}$ denote the entire set of observed entrepreneur-firm combinations, ij, where i indexes entrepreneurs and j indexes firms. Define $\mathcal{S}^1(g) \subset \mathcal{S}$ as the set of entrepreneur-firms ij which are "treated" in year g. For each $ij \in \mathcal{S}^1(g)$, we require each of the following to hold:

1. Entrepreneur i's first child was born in year g.
2. i owns a (weak) plurality of shares in firm j in year $g-1$, and i's ownership share must be no less than $1/3$ in this period.
3. Firm j must have been started in year $g-2$ or earlier.
4. Firm j must not appear in $i'j \in S^1(t)$ for any i' for any $t < g$.

The first requirement is the essence of treatment. The second eliminates minority stakeholders who are more plausibly investors than entrepreneurs. The third ensures that we focus on firms started prior to pregnancy. The fourth removes firms that appeared in the treated sample in an earlier cohort, which might occur if, for example, two individuals co-own a firm and have children in different years.[13]

We define $\mathcal{S}^0(g) \subset \mathcal{S}$ as the set of entrepreneur-firms which are "control" entrepreneur-firms in year g. For each $ij \in \mathcal{S}^0(g)$, we require each of the following to hold:

1. Entrepreneur i did not have their first child $t \in [g-10, 2018]$. 2018 is the final year we observe births.
2. Entrepreneur i owns a (weak) plurality of shares in firm j in year $g-1$, and i's ownership share must be no less than $1/3$ in this period.
3. Firm j must have been started in year $g-2$ or earlier.

[13] We place an additional requirement on male-owned firms, which is that the owner's female spouse must not co-own the firm or work at the firm in $t \in \{-2, -1, 0\}$. This eases the interpretation of our estimates for new fathers and does not substantively change our estimates.

4. Entrepreneur i has no co-owner i' of firm j such that $i'j \in \mathcal{S}^1(t)$ for any t. That is, firm j is not treated via a different co-owner.

The first and fourth requirements isolate a "clean" comparison to entrepreneurs who did not have any children during the comparison period. The second and third requirements are the same as the requirements for the treated sample.

We construct estimation weights so that the weighted joint distribution of firm age and sector in each control sample $\mathcal{S}^0(g)$ matches the distribution in the corresponding treated sample $\mathcal{S}^1(g)$. For an entrepreneur-firm $ij \in \mathcal{S}^0(g) \cup \mathcal{S}^1(g)$, these weights are

$$\omega_{ij}^g = \begin{cases} \theta_{ij}^g & \text{if } i \in \mathcal{S}^0(g) \\ \theta_{ij}^g \left(\sum_{k\ell \in \mathcal{S}^1(g)} \theta_{k\ell}^g \mathbb{1}\{X_{k\ell} = X_{ij}\}\right) \left(\sum_{k\ell \in \mathcal{S}^0(g)} \theta_{k\ell}^g \mathbb{1}\{X_{k\ell} = X_{ij}\}\right)^{-1} & \text{if } i \in \mathcal{S}^1(g) \end{cases}$$
$$\text{with} \quad \theta_{ij}^g = \left(\sum_j \mathbb{1}\{i \text{ owns } j \text{ in } g-1\}\right)^{-1}. \tag{4}$$

The covariates X include firm age (top-coded at 20) and firm sector (SIC section).[14] When we estimate heterogeneity as in Figure 2 and Table 3, we also include the stratifying variable in our matching variables X. The baseline weight θ_{ij}^g is the inverse of the number of firms owned by entrepreneur i in year $g-1$. This ensures that all treated entrepreneurs receive equal weight in our analysis, as opposed to placing greater weight on individuals who own multiple firms.

This weighting procedure applies the intuition of Heckman, Ichimura and Todd (1997, 1998) as extended by Callaway and Sant'Anna (2021). The insight of Heckman, Ichimura and Todd (1997, 1998) is that if treatment is randomly assigned conditional on covariates X, then matching on these covariates will provide a consistent estimate of the average treatment effect on the treated (ATT). Callaway and Sant'Anna (2021) extend this approach (which they call an "outcome regression" approach) to a dynamic setting with multiple cohorts.

We construct a stacked sample as

$$\mathcal{S}^{\text{stack}} = \{\mathcal{S}^0(2002), \mathcal{S}^1(2002), \ldots, \mathcal{S}^0(2018), \mathcal{S}^1(2018)\}. \tag{5}$$

Some observations will naturally appear in the stacked sample multiple times, which occurs

[14] Our sample construction also implicitly conditions on calendar year.

frequently in the control sample. To account for this when conducting inference, we cluster our standard errors two-way on the entrepreneur i and firm j levels.

Using the stacked sample $\mathcal{S}^{\text{stack}}$, we estimate (1) using the implementation developed by Correia (2016). We aggregate across estimates using (2), treating the shares ω_g as fixed when computing the standard errors. This approach follows closely the "stacked regression" recommended by Baker, Larcker and Wang (2022), with the primary difference being that they discuss weighting by the inverse of the estimated variance while we weight by cohort size.

In Figure 1, we transform $\hat{\beta}_t$ into a measure of percent change by following Kleven, Landais and Søgaard (2019). Specifically, we predict the outcome using the regression coefficients from (1) but with D_{ik} fixed at 0:

$$\tilde{Y}_{ijks} = \sum_g \mathbb{1}[k = g] \left(\hat{\alpha}^g + \sum_{t \neq -2} \hat{\lambda}_t^g \mathbb{1}[s = t] \right) + \hat{f}(X_{ijs}). \tag{6}$$

We then take

$$\hat{\pi}_t = \sum_g \omega_g \hat{\pi}_t^g \tag{7}$$

where $\hat{\pi}_t^g$ is the average prediction $\tilde{Y}_{ijks}$ for the treated individuals in cohort g, $ij \in \mathcal{S}^1(g)$. We report in Figure 1 the estimates $\hat{\psi}_t = \hat{\beta}_t / \hat{\pi}_t$. Standard errors are computed by the delta method, noting that $\hat{\pi}_t^g$ can be written as a weighted average of regression coefficients with weights based on the sample characteristics.

Prior to estimation, we make two adjustments to our outcome variable when the outcome is profits or owner salaries. First, we trim the outcome variable at the 99.5th percentile (profits and salaries) and the 0.5th percentile (profits). This is to mitigate the influence of extreme values, and our estimates are not sensitive to these trimming thresholds. Second, if a firm pauses or ceases activity and therefore does not report a balance sheet, we infer that profits and salaries are equal to zero. This allows us to maintain a balanced set of firms in the post-period.

C Matching occupations to industries

We begin with a set of industries based on the Norwegian Standard Industrial Classification (SIC) system[15] and a set of occupations from the Occupational Information Network (O*NET) 28.0 database.[16] For each SIC industry, we observe a brief industry description. For each O*NET occupation, we observe a title and a more detailed occupational description.

We concatenate occupational titles and descriptions and then match them to industry descriptions based on a semantic similarity measure derived from a pre-trained language model. Specifically, we use a pre-trained SentenceTransformer model (all-MiniLM-L6-v2) to encode both industry descriptions and the combined occupation texts into high-dimensional embeddings that capture semantic meaning. Using these embeddings, we compute cosine similarity scores between each industry's description and all occupation vectors. For each industry, the occupation with the highest cosine similarity is identified as the best match. This process systematically associates each industry with the occupation that most closely aligns with its description in terms of semantic content.

Table A2 shows the matched industry-occupations for the most common industries of female owners who have children. The matched occupations appear to provide a representative occupation for their respective industries. Even if it is likely that the entrepreneur is in a different occupation than the matched occupation, this is only problematic if the matched occupation and the true occupation differ in terms of face-to-face contact. For example, although the owner of a restaurant may not wait tables, it is sufficient that the restaurateur engages with clientele at a similar level. In small firms, as those in our data, this seems likely.

[15] https://www.ssb.no/en/klass/klassifikasjoner/6

[16] https://www.onetcenter.org/database.html

Table A2: Industries matched to O*NET occupations

SIC code	Industry description	Matched O*NET occupation
96.020	Hairdressing and other beauty treatment	39-5012.00: Hairdressers, Hairstylists, and Cosmetologists
68.209	Other letting of real estate	11-9141.00: Property, Real Estate, and Community Association Managers
56.101	Operation of restaurants and cafes	35-3031.00: Waiters and Waitresses
86.909	Other human health activities	21-1094.00: Community Health Workers
70.220	Business and other management consultancy activities	13-1111.00: Management Analysts
47.710	Retail sale of clothing in specialised stores	13-1022.00: Wholesale and Retail Buyers, Except Farm Products
85.510	Sports and recreation education	25-1193.00: Recreation and Fitness Studies Teachers, Postsecondary
75.000	Veterinary activities	31-9096.00: Veterinary Assistants and Laboratory Animal Caretakers
86.230	Dental practice activities	31-9091.00: Dental Assistants
68.100	Buying and selling of own real estate	41-9022.00: Real Estate Sales Agents
74.101	Industrial design, product design and other technical design	27-1021.00: Commercial and Industrial Designers

Notes: Industries and occupations are matched as described in the text. This table presents the 10 most common 5-digit SIC industries based on SIC 2007.