Online Appendix to "Industrial Automation, Earnings Parity, and Intra-Household Labor Supply"

Claudio Costanzo*

September 2026

A Data and measurement

This appendix collects the descriptive material behind Sections 2 and 2.1 of the main text: the comparison of the analysis sample with the broader ATUS population, the diffusion of the U.S. robot stock, the distribution of the local shock, and the earnings profiles and robustness checks of the stylized facts.

Table A1 compares the analysis sample of couples with young children to the general ATUS population; the presence of young children compresses leisure for both spouses and redirects the household's time toward market work and active caregiving (Section 2). Figure A1 plots the U.S. operational robot stock over 1993–2019, whose steep post-2000 segment the diary window covers, and Table A2 summarizes the distribution of the exposure measure across commuting zones together with the automotive employment shares underlying it.

Two exhibits support the construction of the stylized facts. Figure A2 reports realized weekly earnings across the wife's share of potential wages, the construction check discussed in Section 2.1, and Table A3 re-estimates the kinked profiles without the couples at equal imputed wages.

B Relative earnings and family beliefs

This appendix documents that stated beliefs about family roles bend at the same earnings threshold as the time-use profiles of Section 2.1. The evidence comes from the *Family and*

*ECARES, Université Libre de Bruxelles. Contact: claudio.costanzo@ulb.be.

Table A1: Descriptive statistics: ATUS general population

	Females			Males		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Leisure (min/day)	122,710	310.07	208.33	96,658	356.87	235.57
Childcare (min/day)	122,710	43.10	101.62	96,658	22.34	69.34
Market work (min/day)	122,710	126.61	217.06	96,658	191.39	261.12
Age	122,710	48.37	18.15	96,658	46.77	17.44
Hourly wage (USD)	35,769	15.08	9.23	29,080	17.52	10.25
Weekly earnings (USD)	61,698	709.03	513.95	54,518	951.10	596.42
Usual work hours per week	113,670	21.20	20.66	88,373	30.94	22.49
Less than secondary (%)	122,710	14.37	35.08	96,658	15.69	36.37
Secondary (%)	122,710	25.91	43.81	96,658	25.52	43.60
Some college (%)	122,710	27.95	44.87	96,658	25.39	43.52
Tertiary (%)	122,710	31.77	46.56	96,658	33.40	47.17

Notes: "General population" includes all ATUS respondents by sex. Variable definitions are as in Table 1.

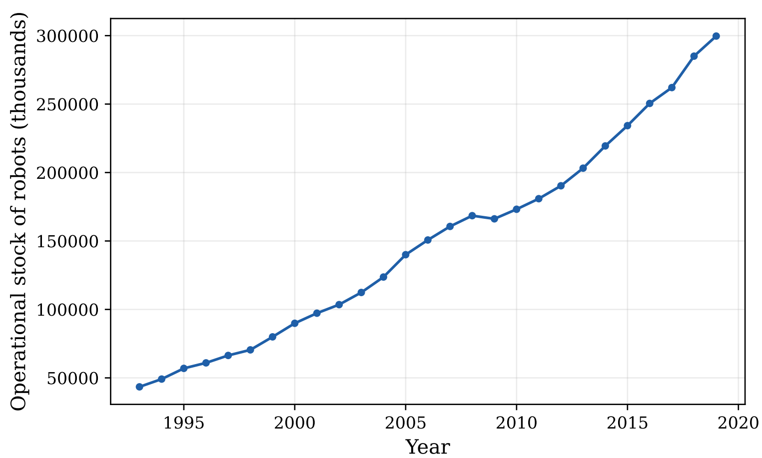

Figure A1: Industrial robot operational stock in the United States, 1993–2019. Data are sourced from the International Federation of Robotics (IFR).

Table A2: Distribution and industrial profile of the robot shock

Statistic	**Value (Robots per 1,000 workers)**
Mean	3.14
Standard deviation	3.25
10th percentile (least affected)	1.29
50th percentile (median)	1.96
90th percentile (most affected)	5.99
99th percentile (extreme)	16.54
Maximum	33.70
Mean automotive share (1990 baseline)	
In least affected areas ($< p10$)	0.62%
In most affected areas ($> p90$)	8.67%

Notes: Statistics are calculated at the Commuting Zone-Year level ($N = 1,859$) for 2003–2019. Least affected examples include New York City; most affected examples include Grand Rapids and Detroit.

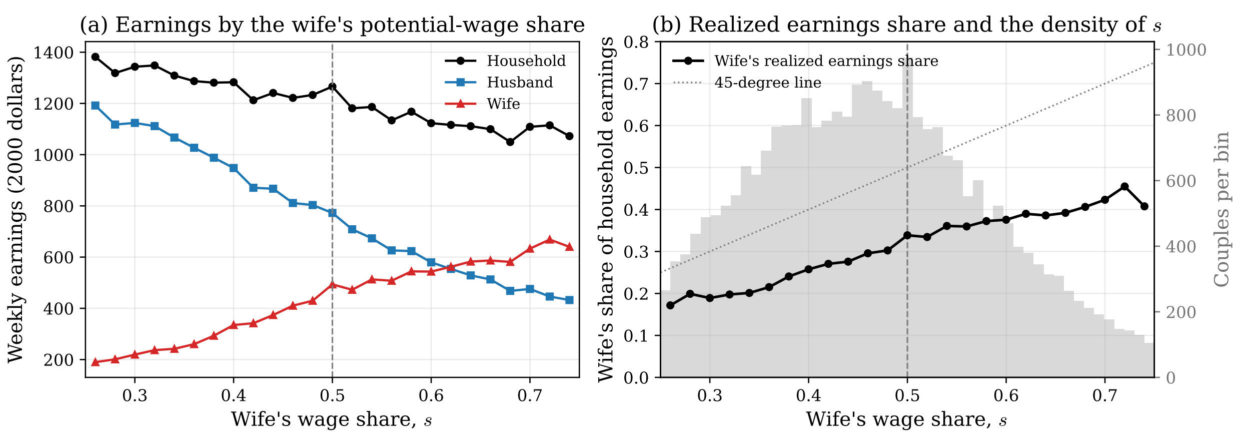

Figure A2: Earnings against the wife's share of potential wages, on the stylized-facts sample of Section 2.1. Panel (a): mean weekly earnings of the household, the husband and the wife (CPS-linked weekly earnings at the main job, zero for the non-employed, constant-2000 dollars; couples with top-coded or unreported earnings dropped) by bins of the wife's share of potential hourly wages s. Panel (b): the wife's realized share of household earnings by bins of s (dots, left axis), the 45-degree line, and the histogram of s (grey, right axis). The realized earnings share rises with s (from 0.25 on average below parity to 0.37 above it) but stays well below s because wives work fewer hours; s has no mass point at 0.5 beyond the 6.6 percent of couples with equal imputed potential wages, whose exclusion leaves the kinked profiles unchanged (Table A3). N = 26,654.

Table A3: The kinked time-use profiles without equal-wage couples

		Baseline s		Excluding $\lvert s-0.5\rvert < 0.01$	
		Slope below	Slope above	Slope below	Slope above
Women	Market work	+181***	−81**	+187***	−87**
		(29)	(39)	(31)	(41)
Women	Childcare	−52***	+98***	−54***	+100***
		(19)	(25)	(20)	(27)
Women	Leisure	−10	−34	−8	−36
		(21)	(28)	(23)	(30)
Men	Market work	+23	+136***	+41	+116**
		(39)	(49)	(42)	(52)
Men	Childcare	+20	−35*	+13	−27
		(16)	(20)	(17)	(21)
Men	Leisure	+26	−103***	+17	−93**
		(28)	(34)	(29)	(36)
Observations (women / men)		14,060 / 12,594		13,135 / 11,757	

Notes: Two-segment linear fits of daily minutes on the wife's wage share s over $s \in (0.25, 0.75)$, joined at parity, on the estimation sample of Section 4 (couples under 64, non-search diary days); slopes in minutes per unit of s with OLS standard errors. *Baseline*: observed hourly wages for workers, cell-imputed potential wages for non-workers (the paper's s). *Excluding*: drops couples whose hourly wages are within one percent of equality, the mass that drives the discontinuity critique of the relative-income literature. The reversals are the sign changes between the two slopes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Changing Gender Roles IV module of the ISSP (ISSP Research Group, 2016), in which respondents report their earnings relative to their spouse on a seven-point scale, from "My spouse/partner has no income" to "I have no income," and state their agreement with a series of family-norm statements, among them "A preschool child is likely to suffer when the mother works" and "Family life suffers when a woman has a full-time job."[1]

Restricting the sample to U.S. couples and excluding the extreme tails of the earnings distribution,[2] I estimate

$$\text{Belief}_i = \alpha + \sum_k \beta_k \, \mathbf{1}(\text{RelEarn}_i = k) + \mu X_i + \varepsilon_i, \tag{A1}$$

where the dependent variable is the agreement score for a given family-norm statement (a higher score indicates disagreement), $\mathbf{1}(\text{RelEarn}_i = k)$ is a vector of dummy variables for the wife's earnings position relative to her husband, and X_i controls for respondent sex, both spouses' ages and education levels, and marriage duration. Standard errors are clustered by respondent and spouse education. Figure A3 plots the resulting marginal effects.

For the items stressing the disadvantages of maternal employment or endorsing traditional spousal roles (Panels a–d), the disagreement score follows an inverse-U: respondents move from agreement when the wife earns far less, to strong disagreement near earnings parity, and back toward agreement once the wife clearly out-earns the husband. The reversion suggests that couples in non-traditional earnings arrangements adopt more traditional stated attitudes to compensate for the deviation from the norm. Attitudes toward maternal employment when children are older (Panel f) instead trace a V in the disagreement score, with couples at parity the most likely to agree that women should work, and the preschool-age item (Panel e) shows a similar but weaker pattern, indicating less sensitivity to relative earnings on that margin. Beliefs thus turn at the threshold where the time-use profiles turn, the correspondence that Section 2.1 cites.

C Design and instrument diagnostics

This appendix reports the diagnostics behind the design of Section 3.1: the first-stage estimates, the alternative EU5 shifters, the Rotemberg and share-balance diagnostics of the shares-based identification, the leave-automotive exercises, the covariate-trend specifications,

[1] Two additional items, "Being a housewife is as fulfilling as working for pay" and "Both partners should contribute to household income," were dropped because their directional interpretation regarding traditionalism is ambiguous.

[2] Including extreme points strengthens the observed dynamics but may confound the results with involuntary unemployment effects.

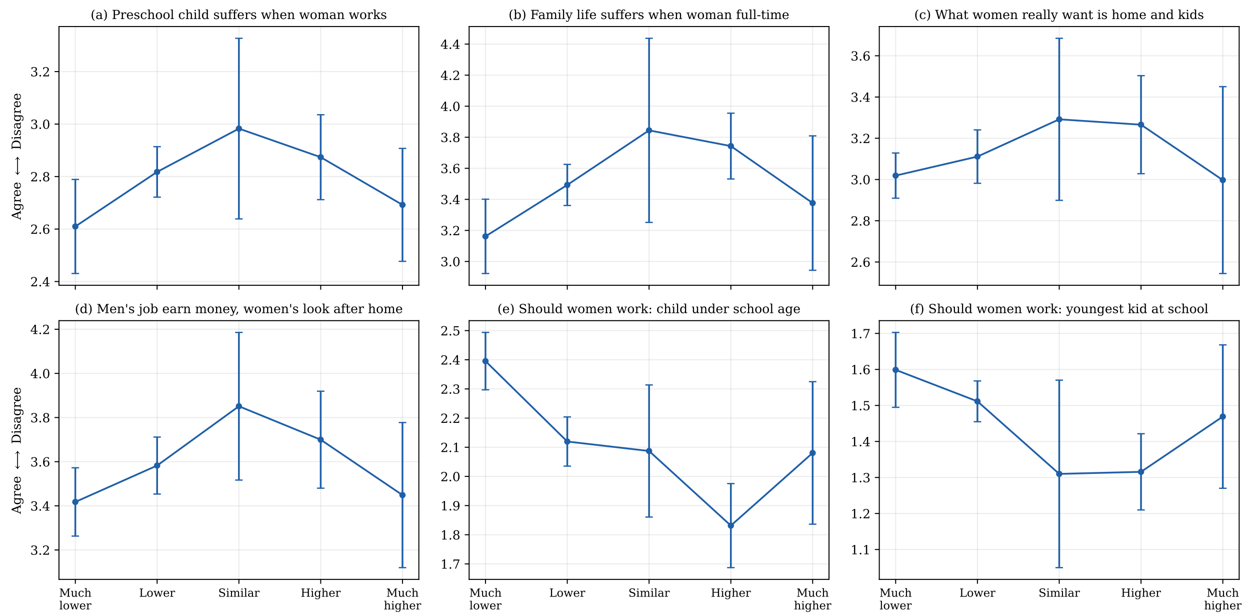

Figure A3: Marginal effects of the wife's relative earnings on agreement with family-norm statements. The y-axis represents the disagreement score (higher = less traditional). The x-axis represents the wife's contribution to household income, ranging from much less than the husband (left) to much more (right).

and the pre-trend placebos.

First stage and alternative shifters. Table A4 reports the first-stage estimates underlying Table 4. Tables A5, A6 and A7 re-estimate the market-work, childcare and leisure responses instrumenting with each EU5 country's series one at a time and with shifters lagged three and five years; all five countries deliver the female withdrawal, with reduced forms that change sign together with the first stage, as Section 3.1 discusses.

Accounting for the automotive industry. Shift-share estimators can be sensitive to a small number of industries that receive disproportionate weight, which compromises the estimator's asymptotic properties; I therefore compute the Rotemberg weights for the baseline Bartik measure. Table A8 shows that the vehicle sector receives a large weight, reflecting its status as a primary adopter of industrial robots. Because the automotive industry constitutes a concentrated shock, Table A9 controls for commuting-zone trends across quartiles of the 1990 vehicle-sector employment share; the point estimates for both men and women remain stable, so the intra-household dynamics are a general response to automation and no artifact of one industry's trends.

As a more demanding check, Table A10 rebuilds both the exposure measure and the EU5 instrument after removing the automotive component entirely. Because the automotive

Table A4: First stage: local robot exposure on the EU5 external shifter.

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Robots$_{ct}$						
EU5 shifter	5.259***	5.208***	5.332***	4.772***	4.891***	4.820***
	(0.280)	(0.267)	(0.278)	(0.236)	(0.242)	(0.264)
KP F	353	380	368	409	408	334
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: First-stage estimates of local robot exposure on the EU5 external shifter, within each estimation subsample of the main table (Table 4). Controls, fixed effects, ATUS weights, and CZ–year clustering exactly as in the 2SLS. The just-identified Kleibergen–Paap F equals the squared cluster-robust t-statistic on the shifter. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

sector is the dominant adopter of industrial robots, it carries most of the identifying variation: the residual instrument is much weaker, so the market-work coefficients become imprecise. The gendered childcare reallocation, however, survives (fathers' childcare falls and mothers' rises, the same specialization pattern as in the baseline), indicating that the qualitative result, a household move away from the added-worker response and toward a traditional division of labor, is no artifact of the automotive sector alone. The result states where the identifying variation lives: the natural experiment is the diffusion of industrial robots, which is concentrated in automotive, and the time-use response to it is gendered in the direction the breadwinner-norm mechanism predicts. The full Rotemberg diagnostics (Goldsmith-Pinkham et al., 2020) in Table A11 make the same point quantitatively: automotive carries 56% of the Rotemberg weight and its own just-identified estimate (−13.1 minutes, first-stage $F = 264$) coincides with the headline, the sum of negative weights is negligible (−0.011), and the effective number of independent shocks is small (about three). The design is thus best understood as identified off the diffusion of automotive robots, the principal site of industrial automation, and it should be interpreted accordingly.

Share exogeneity and the cultural confound. A shift-share design built on historical industry shares requires that those shares be uncorrelated with confounding regional characteristics (Goldsmith-Pinkham et al., 2020). Table A12 reports the cross-commuting-zone correlations of the high-weight 1970 automotive share, and of the EU5 instrument, with predetermined 1990 characteristics. The shares load on manufacturing intensity, as expected for a Bartik measure, but the most consequential entry for the present paper is the *negative* correlation between exposure and traditional religious adherence (−0.18 for the automotive

Table A5: Alternative EU5 shifters: lagged and single-country instruments, market work (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Market work						
EU5 mean (baseline)	12.734**	8.967**	13.953**	-15.949***	-18.510***	-16.361***
	(5.834)	(3.941)	(6.306)	(4.767)	(5.749)	(6.264)
	+5.26 [353]	+5.21 [380]	+5.33 [368]	+4.77 [409]	+4.89 [408]	+4.82 [334]
EU5 mean, lagged 3 years	17.686***	11.168***	18.973***	-15.906***	-15.026**	-15.793**
	(4.615)	(3.500)	(4.899)	(4.495)	(6.183)	(6.678)
	+4.79 [642]	+4.76 [693]	+4.83 [653]	+4.62 [450]	+4.76 [429]	+4.75 [412]
EU5 mean, lagged 5 years	15.640***	9.861***	14.966***	-14.748***	-13.059**	-14.389**
	(4.143)	(3.192)	(4.549)	(4.059)	(5.517)	(5.910)
	+3.84 [1099]	+3.79 [1143]	+3.87 [1166]	+3.82 [637]	+3.89 [476]	+3.88 [552]
Denmark only	14.080***	9.547***	14.361***	-17.004***	-13.633***	-13.656**
	(4.669)	(3.210)	(5.219)	(3.820)	(4.872)	(5.333)
	+2.51 [1295]	+2.49 [1199]	+2.52 [1251]	+2.47 [1534]	+2.50 [1222]	+2.50 [1231]
Finland only	9.281	6.492*	6.939	-18.277***	-19.878***	-17.528***
	(6.817)	(3.569)	(7.462)	(4.272)	(5.274)	(5.859)
	+1.16 [156]	+1.15 [200]	+1.16 [151]	+1.09 [308]	+1.08 [333]	+1.07 [285]
France only	11.542	11.238**	7.287	-20.384***	-20.598***	-19.783***
	(7.346)	(4.394)	(7.825)	(4.993)	(5.421)	(6.189)
	−2.90 [141]	−2.83 [201]	−2.90 [138]	−2.77 [213]	−2.76 [247]	−2.76 [212]
Italy only	6.792	-0.493	2.026	-17.747***	-14.899**	-12.704*
	(8.439)	(5.580)	(8.971)	(5.627)	(6.399)	(6.995)
	−1.72 [102]	−1.69 [143]	−1.73 [98]	−1.52 [122]	−1.54 [126]	−1.57 [126]
Sweden only	11.690*	4.548	12.109*	-13.860***	-19.322***	-15.358***
	(6.662)	(4.489)	(6.926)	(5.028)	(5.351)	(5.781)
	+3.66 [189]	+3.56 [250]	+3.76 [193]	+3.17 [301]	+3.19 [283]	+3.21 [270]
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering) and the baseline estimation sample, replacing the instrument. The baseline shifter averages the 1970-share-weighted penetration of Denmark, Finland, France, Italy and Sweden; the lagged rows use the EU5 average measured three and five years before the diary year (rebuilt from the 1970 shares and the IFR EU5 panel), and the single-country rows use one national series alone. The third line of each row reports the first-stage coefficient on the instrument and, in brackets, the Kleibergen–Paap first-stage F. The French and Italian series enter the first stage with a negative sign because their automotive penetration declined over 2004–2019 while U.S. penetration rose, so that in automotive-specialized zones these series fall relative to other zones as U.S. exposure rises; leave-one-country-out averages (replication files) reproduce the baseline except that the average without Finland, in which the positive Nordic and negative French–Italian components offset, has a first-stage F below ten. Observations as in Table 4. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6: Alternative EU5 shifters: lagged and single-country instruments, childcare (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Childcare						
EU5 mean (baseline)	0.854	2.466	1.771	6.579**	9.686***	10.092***
	(2.260)	(2.031)	(2.214)	(2.627)	(2.815)	(3.030)
	+5.26 [353]	+5.21 [380]	+5.33 [368]	+4.77 [409]	+4.89 [408]	+4.82 [334]
EU5 mean, lagged 3 years	-0.230	1.497	-0.378	2.872	6.176**	7.103**
	(1.886)	(1.751)	(1.899)	(2.509)	(2.952)	(3.154)
	+4.79 [642]	+4.76 [693]	+4.83 [653]	+4.62 [450]	+4.76 [429]	+4.75 [412]
EU5 mean, lagged 5 years	0.564	1.630	0.673	1.439	5.385**	6.268**
	(1.717)	(1.628)	(1.735)	(2.309)	(2.568)	(2.763)
	+3.84 [1099]	+3.79 [1143]	+3.87 [1166]	+3.82 [637]	+3.89 [476]	+3.88 [552]
Denmark only	2.481	1.507	2.627	1.537	2.460	3.321
	(1.741)	(1.633)	(1.777)	(2.409)	(2.821)	(3.019)
	+2.51 [1295]	+2.49 [1199]	+2.52 [1251]	+2.47 [1534]	+2.50 [1222]	+2.50 [1231]
Finland only	3.498	2.954*	4.593**	3.514	7.319***	8.320***
	(2.182)	(1.611)	(2.184)	(2.323)	(2.027)	(2.305)
	+1.16 [156]	+1.15 [200]	+1.16 [151]	+1.09 [308]	+1.08 [333]	+1.07 [285]
France only	3.553	2.595	4.034	1.588	6.836**	7.983**
	(2.367)	(2.020)	(2.465)	(2.973)	(2.759)	(3.154)
	−2.90 [141]	−2.83 [201]	−2.90 [138]	−2.77 [213]	−2.76 [247]	−2.76 [212]
Italy only	6.962**	3.007	7.720***	-4.803	-3.559	-2.682
	(2.908)	(2.497)	(2.947)	(4.916)	(5.804)	(6.291)
	−1.72 [102]	−1.69 [143]	−1.73 [98]	−1.52 [122]	−1.54 [126]	−1.57 [126]
Sweden only	0.582	3.720*	1.706	3.800	7.310**	6.578*
	(2.571)	(2.049)	(2.535)	(2.985)	(3.230)	(3.614)
	+3.66 [189]	+3.56 [250]	+3.76 [193]	+3.17 [301]	+3.19 [283]	+3.21 [270]

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering) and the baseline estimation sample, replacing the instrument. The baseline shifter averages the 1970-share-weighted penetration of Denmark, Finland, France, Italy and Sweden; the lagged rows use the EU5 average measured three and five years before the diary year (rebuilt from the 1970 shares and the IFR EU5 panel), and the single-country rows use one national series alone. The third line of each row reports the first-stage coefficient on the instrument and, in brackets, the Kleibergen–Paap first-stage F. Observations as in Table A5. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A7: Alternative EU5 shifters: lagged and single-country instruments, leisure (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Leisure						
EU5 mean (baseline)	-5.494	-3.772	-8.935***	6.205**	10.996***	9.746***
	(3.930)	(3.184)	(3.230)	(3.057)	(3.062)	(3.659)
	+5.26 [353]	+5.21 [380]	+5.33 [368]	+4.77 [409]	+4.89 [408]	+4.82 [334]
EU5 mean, lagged 3 years	-5.724	-3.268	-8.625***	8.579***	10.538***	11.730***
	(4.326)	(3.017)	(3.307)	(3.095)	(2.986)	(3.397)
	+4.79 [642]	+4.76 [693]	+4.83 [653]	+4.62 [450]	+4.76 [429]	+4.75 [412]
EU5 mean, lagged 5 years	-6.829*	-4.264	-8.860***	7.587***	8.926***	10.707***
	(4.134)	(2.873)	(3.090)	(2.918)	(3.014)	(3.350)
	+3.84 [1099]	+3.79 [1143]	+3.87 [1166]	+3.82 [637]	+3.89 [476]	+3.88 [552]
Denmark only	-7.206*	-4.913*	-9.193***	8.060***	10.588***	11.652***
	(4.135)	(2.727)	(3.190)	(2.739)	(2.812)	(3.373)
	+2.51 [1295]	+2.49 [1199]	+2.52 [1251]	+2.47 [1534]	+2.50 [1222]	+2.50 [1231]
Finland only	-4.030	-5.210*	-6.009**	6.433**	11.618***	10.456***
	(3.570)	(2.948)	(3.039)	(2.803)	(3.308)	(3.817)
	+1.16 [156]	+1.15 [200]	+1.16 [151]	+1.09 [308]	+1.08 [333]	+1.07 [285]
France only	-1.348	-7.049**	-2.038	7.451**	12.696***	12.348***
	(4.356)	(3.286)	(3.832)	(3.046)	(3.574)	(3.958)
	−2.90 [141]	−2.83 [201]	−2.90 [138]	−2.77 [213]	−2.76 [247]	−2.76 [212]
Italy only	-5.611	-5.116	-5.428	9.963***	12.991***	13.172***
	(4.792)	(3.845)	(4.394)	(3.629)	(4.051)	(4.534)
	−1.72 [102]	−1.69 [143]	−1.73 [98]	−1.52 [122]	−1.54 [126]	−1.57 [126]
Sweden only	-2.655	-2.307	-5.680	8.738***	14.198***	11.705***
	(4.004)	(3.311)	(3.695)	(3.029)	(3.273)	(3.528)
	+3.66 [189]	+3.56 [250]	+3.76 [193]	+3.17 [301]	+3.19 [283]	+3.21 [270]

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering) and the baseline estimation sample, replacing the instrument. The baseline shifter averages the 1970-share-weighted penetration of Denmark, Finland, France, Italy and Sweden; the lagged rows use the EU5 average measured three and five years before the diary year (rebuilt from the 1970 shares and the IFR EU5 panel), and the single-country rows use one national series alone. The third line of each row reports the first-stage coefficient on the instrument and, in brackets, the Kleibergen–Paap first-stage F. Observations as in Table A5. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A8: Rotemberg weights by industry

Industry	**Weight**	**Industry**	**Weight**
Automotive	0.563	Furniture	0.002
Manufacturing, other	0.130	Paper	0.001
Electronics	0.091	Utilities	0.001
Metal Basic	0.083	Textiles	0.001
Petrochemicals	0.050	Research	0.001
Food	0.037	Mining	0.000
Metal Products	0.019	Agriculture	0.000
Metal Machinery	0.017	Construction	0.000
Mineral	0.010	Services	−0.011
Vehicles Other	0.005		

Notes: Rotemberg (Goldsmith-Pinkham-Sorkin-Swift) weights α_k, normalized to sum to one with negative weights preserved (matching the diagnostics in Table A11). Automotive carries 0.563 of the weight; the sum of the negative weights is −0.011.

share, −0.20 for the instrument). The automotive belt is *less* religiously traditional than the rest of the country: the manufacturing heartland and the culturally conservative South and Intermountain West are distinct regions. This directly addresses the concern that the amplified response in conservative areas (Table 7) merely reflects those areas' industrial structure. If conservative commuting zones were the high-exposure auto belt, the interaction would be confounded; they are not, so the cultural heterogeneity is identified off variation that is orthogonal to, and in fact opposes, the geography of the shock.

Covariate-specific trends. A share-based design can be confounded if the zones that carry the identifying shares were on their own trajectories for reasons unrelated to robots. Table A13 adds to the baseline the interactions of six predetermined 1990 characteristics (manufacturing and female manufacturing employment shares, college share, female population share, log population, and the share of traditional religious adherents, each standardized) with a linear trend and, separately, with year effects; state-by-year effects remain in every specification, so the interactions absorb within-state differential trends. The female market-work withdrawal is robust to both sets of trends and larger than in the baseline (−18.6 to −24.2 minutes with linear trends, −17.8 to −19.8 with year effects); the childcare increase survives the linear trends among employed and dual-earner mothers (+9.9 and +9.6) but not the year effects, and the leisure increase weakens. The male market-work increase does not survive either set: it falls to +2.1, +0.3 and +2.1 minutes with linear trends and to +1.1, −1.6 and −0.7 with year effects, while the first-stage F remains between 260 and 420. The pattern is not driven by any single characteristic: adding one trend at a time leaves the male estimates at

Table A9: Effect of robots on time use, adjusting for commuting-zone-specific trends across quartiles of employment share in the automotive sector (2SLS).

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	13.010**	10.723***	14.277**	-13.619***	-14.210**	-13.426**
	(6.025)	(3.646)	(6.517)	(4.826)	(5.982)	(6.477)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.436	0.450	0.486	0.292	0.431	0.440
Daily minutes childcare						
Robots	1.891	2.449	2.437	6.320**	10.997***	12.403***
	(2.292)	(1.769)	(2.252)	(2.864)	(2.995)	(3.545)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.243	0.193	0.268	0.249	0.246	0.265
Daily minutes leisure						
Robots	-4.528	-4.486	-8.809***	6.195*	11.045***	8.882**
	(4.042)	(3.262)	(3.250)	(3.181)	(3.364)	(3.869)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.302	0.266	0.324	0.213	0.311	0.320
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Sample includes married or cohabiting couples with at least one spouse working and children aged 10 or less. Columns 1 and 4 capture the extensive margin of the respondent (spouse is employed). Columns 2 and 5 capture the extensive margin of the spouse (respondent is employed). Columns 3 and 6 restrict to dual-earner households, capturing only the intensive margin of labor supply. Control variables include commuting zone and state-by-year FE, household demographics, day/month/holiday. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10: Robustness: robot instrument excluding the automotive sector (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots (no auto)	44.820	90.042	70.383	-3.391	-35.791	-20.432
	(51.193)	(65.963)	(51.267)	(33.565)	(35.728)	(38.944)
Daily minutes childcare						
Robots (no auto)	-39.519*	-20.528	-37.770*	47.659*	43.762	37.663
	(21.942)	(29.870)	(21.916)	(28.185)	(31.064)	(35.718)
Daily minutes leisure						
Robots (no auto)	17.198	21.694	7.010	-3.201	22.388	14.685
	(32.742)	(34.270)	(32.097)	(21.099)	(21.336)	(24.629)
Observations	4,506	7,306	4,193	7,967	5,400	4,844

Notes: Main 2SLS with both robot exposure and the EU5 instrument rebuilt after dropping the automotive exposure component (the largest Rotemberg weight). The first stage is much weaker without automotive, so market-work estimates are imprecise; the gendered childcare reallocation (fathers' care down, mothers' up) nonetheless persists. Controls and FE as in the main table; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A11: Rotemberg diagnostics (Andrews–Gentzkow–Shapiro): top industries, women's market work

Industry	α_k	β_k	F_k
Automotive	$+0.563$	-13.10**	264
Manufacturing (other)	$+0.130$	weak IV	1
Electronics	$+0.091$	$+3.48$	27
Basic metals	$+0.082$	weak IV	6
Petrochemicals	$+0.049$	weak IV	0
Food	$+0.037$	weak IV	7

Notes: For each high-weight industry: the Rotemberg weight α_k (normalized to sum to one, negatives preserved), the just-identified 2SLS estimate β_k using only that industry's EU5 shock (women, dual-earner market work), and its first-stage F_k ("weak IV" marks $F_k < 10$). Sum of negative weights $= -0.011$; effective number of shocks (inverse Herfindahl) = 2.8. Automotive carries the identifying variation and its own estimate coincides with the headline; consistent with the leave-automotive-out check (Table A10).

Table A12: Share-exogeneity / pre-period balance of the robot instrument

1990 commuting-zone characteristic	1970 automotive share	EU5 instrument
Female share	+0.144 (0.037)	+0.213 (0.047)
College share	−0.159 (0.059)	−0.178 (0.062)
High-school share	+0.141 (0.056)	+0.218 (0.060)
Manufacturing share	+0.329 (0.055)	+0.484 (0.052)
Log population	+0.161 (0.076)	+0.202 (0.067)
Black share	+0.011 (0.059)	−0.003 (0.061)
Traditional adherence	**−0.175** (0.030)	**−0.204** (0.037)

Notes: Cross-commuting-zone correlations of the high-weight 1970 automotive employment share (col. 1) and the EU5 instrument (col. 2) with predetermined 1990 characteristics; bootstrap standard errors over commuting zones (2,000 draws) in parentheses. The negative traditional-adherence correlation shows the automotive belt is not the religiously conservative belt, so the norm-heterogeneity interaction of Table 7 is not confounded by industrial structure. CZ-level, unweighted.

10–13 minutes, significant in most cases, and the set without the overall manufacturing-share trend gives 10 minutes at the ten-percent level; it is the two manufacturing-share trends together that absorb the male response. Because exposure is a manufacturing-share-weighted national trend, the residual variation once manufacturing-specific trends are removed is the part of the automotive share orthogonal to overall manufacturing intensity, together with the non-linearity of the robot stock's path, which evidently suffices for the female response but not for the male one. The male intensive-margin estimate is accordingly read as less securely identified than the female withdrawal.

Pre-trends. A remaining concern is that commuting zones destined for heavy robot exposure were already on distinct gender-gap trajectories. Because industrial robots diffused overwhelmingly after 2000, the change in the commuting-zone gender gap over the *preceding* decade (1990–2000) is a placebo for the design. Regressing the 1990–2000 change in the gender employment gap, and in the gender wage gap, on the later robot-exposure change (with the same 1990 covariates and Census-division fixed effects, on 1990 and 2000 decennial Census microdata) yields coefficients that are economically negligible and statistically indistinguishable from zero (employment gap +0.001, $t = 0.4$; wage gap +0.001, $t = 0.5$). Running the same placebo on participation *levels* instead of gaps shows why: labor-force participation was rising in soon-to-be-exposed zones over the 1990s for *both* sexes, and by almost exactly the same amount (+1.2 percentage points per unit of later exposure for women, $t = 5.5$; +1.3 for men, $t = 5.2$), the manufacturing-belt boom of the pre-robot decade. This common level trajectory differences out of every gendered contrast the paper uses, which is why the gap placebos are null; and for the female-work outcome specifically it runs *opposite* in sign

Table A13: Trend-augmented specifications: 1990 commuting-zone characteristics interacted with time (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Market work						
Baseline	12.734**	8.967**	13.953**	-15.949***	-18.510***	-16.361***
	(5.834)	(3.941)	(6.306)	(4.767)	(5.749)	(6.264)
1990 characteristics × trend	2.097	0.247	2.105	-18.621***	-24.169***	-22.073***
	(7.547)	(5.328)	(8.072)	(5.730)	(7.374)	(7.900)
1990 characteristics × year effects	1.144	-1.640	-0.646	-19.841***	-18.326***	-17.808**
	(6.549)	(4.679)	(6.953)	(5.488)	(7.001)	(7.419)
First-stage F	400	377	421	430	489	460
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Childcare						
Baseline	0.854	2.466	1.771	6.579**	9.686***	10.092***
	(2.260)	(2.031)	(2.214)	(2.627)	(2.815)	(3.030)
1990 characteristics × trend	0.369	3.281	2.204	4.930	9.859***	9.546***
	(3.077)	(2.701)	(2.965)	(3.097)	(3.118)	(3.662)
1990 characteristics × year effects	-0.745	3.302	0.914	3.734	5.798	5.146
	(3.434)	(2.342)	(3.314)	(3.454)	(3.580)	(4.609)
First-stage F	400	377	421	430	489	460
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Leisure						
Baseline	-5.494	-3.772	-8.935***	6.205**	10.996***	9.746***
	(3.930)	(3.184)	(3.230)	(3.057)	(3.062)	(3.659)
1990 characteristics × trend	1.661	2.334	-2.122	3.290	9.979**	7.616
	(5.068)	(4.406)	(4.304)	(3.652)	(3.983)	(4.651)
1990 characteristics × year effects	3.235	5.257	1.995	2.005	6.872*	3.644
	(4.449)	(3.769)	(4.357)	(3.723)	(3.866)	(4.444)
First-stage F	400	377	421	430	489	460
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering), on the observations with non-missing 1990 characteristics. The augmented rows add the interactions of six predetermined 1990 commuting-zone characteristics (manufacturing employment share, female manufacturing employment share, college share, female population share, log population, and the share of traditional religious adherents; each standardized) with a linear trend or with year effects; state-by-year effects are in every row, so the interactions absorb within-state differential trends of zones that differ in industrial structure, education, size or norms. First-stage F refers to the year-effects row. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to the estimated effect, so a continuation of pre-period trends would bias the design away from finding a female withdrawal. Finally, the couple-level margin itself shows no differential pre-trend: the 1990–2000 change in a zone's near-parity couple share is unrelated to later exposure ($t = 0.2$), and the change in the mean wife's wage share is small and at most marginal (+0.001, $t = 1.7$). Areas that would later be heavily exposed were not on differential gendered trajectories, consistent with the identifying assumption that the post-2000 robot wave is the source of the variation.

D Robustness checks in detail

This appendix discusses in full the robustness checks summarized in Sections 3.2 and 3.4 of the main text, together with the supporting evidence for the mechanism subsection (Section 3.5); each table sits with the paragraph that discusses it.

The wage first stage in the ATUS sample. The ATUS is small relative to the Census or the ACS, so a first check is whether the labor-market effects of robotization estimated inside the ATUS sample look like the ones the literature documents at scale; if they did not, the time-use results might say more about the sample than about the economy. I therefore estimate commuting-zone-level wage regressions on aggregated ATUS microdata. Wages call for a different design than the diaries. Local labor markets absorb structural shocks over the medium run (Acemoglu and Restrepo, 2020), and annual commuting-zone ATUS cells hold few respondents, so annual two-way fixed effects would mostly difference sampling noise. I therefore use stacked long differences over 3, 5, and 7 years, which smooth the high-frequency noise and capture the medium-run adjustment.

Table A14 shows that the ATUS sample captures the same wage erosion the Census-based studies document. At the 3-year difference, a unit of exposure lowers men's hourly wages by 6.1 percent and narrows the log gender wage gap by 4.9 points ($\beta = -0.061$ and -0.049, Columns 1 and 3); the male effect persists, attenuated, at 5 and 7 years. Within the interquartile range of the shock this is an erosion of several percent of the primary earner's potential wage, the income shock against which the wives' withdrawal must be read, where an added-worker expansion would have been the textbook response.

These wage results are also the *first stage* for the paper's mechanism: by depressing men's potential wages while leaving women's roughly unchanged, robot exposure raises the wife's relative potential wage and narrows the within-couple wage gap (Table A14), pushing the distribution of household wage shares toward parity. Estimated directly on observed dual-earner couples, the effect of exposure on the wife's wage share is a precise zero

Table A14: Effect of robot exposure on wages and wage gap (3, 5, and 7-year differences, 2SLS).

	Male Wages (1)	**Female Wages** (2)	**Wage Gap** (3)
Panel A: 3-Year Differences			
Robots	-0.061***	-0.010	-0.049***
	(0.011)	(0.010)	(0.014)
Observations	1,148	1,184	1,068
R^2	0.031	0.014	0.024
Panel B: 5-Year Differences			
Robots	-0.024**	0.004	-0.026
	(0.010)	(0.012)	(0.018)
Observations	929	949	857
R^2	0.027	0.014	0.011
Panel C: 7-Year Differences			
Robots	-0.014**	0.003	-0.016*
	(0.006)	(0.008)	(0.009)
Observations	734	754	685
R^2	0.059	0.012	0.029

Notes: Sample: Working age individuals (25–54) in the ATUS. The dependent variable is the long difference (3, 5, or 7 years) of the log hourly wage for men, women, or the gender gap. All specifications include 1990 baseline controls and year fixed effects. Robust standard errors clustered at the commuting-zone level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(Table A15); so is the effect on the *potential*-wage share computed over all couples with the same state-inclusive imputation the structural model uses. Translating the latter estimate into flows, the share of couples the shock mechanically pushes across the parity threshold is essentially nil. For the observed share the null is partly what the endogeneity of realized wages predicts: a wife who withdraws as she nears parity lowers her own measured earnings and hours, the simultaneity that has long bedeviled this literature (Heggeness and Murray-Close, 2019; Rosenberg, 2021). For the potential share that simultaneity does not apply. To establish whether the ATUS null reflects limited power or a small movement, I rebuild the couple-level first stage at Census scale: 2.30 million married couples constructed from ACS records with spouse-attached characteristics, collapsed to commuting zones and long-differenced over the same window as Table 3 (Table A16). The estimates settle both sign and magnitude. Without controls, exposure moves couples toward parity in the direction the mechanism requires: the wife's potential-wage share rises, and the share of wives out-earning their husbands rises by 0.13 percentage points per robot, significant at the one-percent level at this scale, with the same pattern among couples with young children. The couple-level movement is therefore real, and the ATUS null was partly a power problem. But the magnitude is much too small for within-couple parity-crossing to carry the time-use response: a three-robot shock shifts roughly 0.4 percent of couples across the threshold, and 0.04 percent once 1990 covariates and division fixed effects absorb regional trends. At this sample size the controlled estimate is precisely zero, with ample power to detect a larger movement. The relative-wage channel is thus established at the labor-market level, and it nudges the couple distribution toward parity, but the reduced-form time-use response cannot be transmitted couple-by-couple through mechanical threshold crossings; this is the same conclusion the structural decomposition reaches from the model side (Section 4), where even the calibrated shock's 8.7% crossing rate could not generate the observed withdrawal. The response is better read as households reallocating when the local labor market moves against the male-breadwinner occupation mix, with its intensity governed by the local salience of the norm (Table 7) and the couple's position in the wage-share distribution (Section 2.1), than as a per-couple parity event. The structural model is accordingly identified off the cross-sectional time-use profile in s; a first stage of the shock on s plays no role. The model contributes a restriction on mechanisms: only an identity wedge reproduces the childcare reversals (Section 4.4). It does not supply a second, independent causal identification of the mediator. I return to the direct, reduced-form test of whether the response concentrates near parity below.

Table A15: Effect of robot exposure on the wife's relative wage (2SLS)

	Effect on share s	Effect on $\mathbf{1}[s > 0.5]$	Observations
Observed (dual-earner)	−0.001 (0.003)	−0.004 (0.012)	6,538
Potential (all couples)	−0.000 (0.002)	+0.006 (0.007)	14,912

Notes: 2SLS of the wife's wage share $s = w_f/(w_f + w_m)$ and an indicator that she out-earns her husband on the hourly rate, on robot exposure instrumented by the EU5 shifter. "Observed" uses measured hourly wages among dual-earner couples; "Potential" imputes non-workers' wages from sex×education×age×state×year cells (with coarser fallbacks), covering all couples and matching the structural model's state-inclusive s. The two readings differ. For the *observed* panel the zero partly reflects the endogeneity/selection of realized wages discussed in Section 3.4 (a wife who withdraws lowers her own measured share). For the *potential* panel that simultaneity does not apply, yet the coefficient is still a precise zero *even with* the geographically varying state-inclusive imputation: robot exposure does not detectably move any individual couple's relative *potential* wage in the ATUS, even though it narrows the gender wage gap at the local-labor-market level (Table A14, Table 3). The couple-level magnitude of the relative-wage shift is thus too small to identify in the ATUS micro data; the relative-wage channel is established at the labor-market level and cannot be traced couple by couple, and the structural model is identified off the cross-sectional profile; this first stage plays no role in it (see Section 3.4). Controls and FE as in the main table; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Census-scale corroboration of the first stage. To confirm that this labor-market first stage is not specific to the modest ATUS sample, I replicate it in the American Community Survey, roughly eighteen million working-age adults. I form a long difference between a

Table A16: Couple-level wage-share first stage at Census scale (ACS couples, 2SLS long difference)

	No controls		+ 1990 controls, division FE	
Outcome (long diff 2005–07 → 2015–17)	Coef.	(SE)	Coef.	(SE)
Wife's wage share, observed (dual-earner)	+0.00037**	(0.00015)	+0.00000	(0.00018)
Wife's wage share, potential (all couples)	+0.00044***	(0.00014)	−0.00004	(0.00014)
Share of wives out-earning (potential)	+0.00134***	(0.00049)	+0.00012	(0.00047)
Share of couples near parity (potential)	+0.00137***	(0.00047)	+0.00019	(0.00042)
Potential share, young-child couples	+0.00049***	(0.00016)	−0.00003	(0.00021)
Out-earning share, young-child couples	+0.00140***	(0.00052)	+0.00026	(0.00072)

Notes: Couple-level analogue of Table 3: 2.30 million married mixed-sex couples (wife aged 24–55) built from ACS records with spouse-attached characteristics (Ruggles et al., 2024), collapsed to commuting-zone means and long-differenced. The wife's wage share is her share of the couple's hourly wages: "observed" uses reported wages of dual-earner couples; "potential" imputes non-earners from employed cell means (sex × education × five-year age × state × year, broad-cell fallback), mirroring the paper's imputation. "Near parity" is a potential share within 0.1 of one half; the young-child rows restrict to couples whose youngest child is ten or younger (the ATUS mirror). 2SLS with the EU5 instrument; right panel adds 1990 covariates and Census-division fixed effects; weighted by base-period couple population; heteroskedasticity-robust standard errors. Without controls, exposure nudges couples toward parity with the sign the mechanism requires; the magnitudes are nonetheless tiny (a three-robot shock moves the out-earning share by +0.4 percentage points) and vanish with the controls, so mass within-couple parity-crossing is not the transmission channel (see text). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

pre-recession base (2005–2007 pooled) and a recovery period (2015–2017 pooled), map ACS respondents to commuting zones with population crosswalks, and regress the change in each zone's mean log hourly wage and employment rate, by sex, on the change in robot exposure (instrumented by the EU5 shifter), using the same 1990 covariates and Census-division fixed effects as the ATUS long difference. Table 3 in the main text reports the estimates. Robot exposure significantly erodes men's hourly wages (−0.0036), and by more than women's (−0.0027), so the wife's relative wage rises, the same movement the ATUS estimates identify, now in a sample two orders of magnitude larger. The estimates echo the paper's mechanism in two respects. The wage response loads on the male side, consistent with automation devaluing male-dominated routine tasks. Men's *employment*, meanwhile, does not fall; the point estimate *rises* with exposure (+0.0020). This estimate should be read with caution. It runs opposite to the net displacement documented by Acemoglu and Restrepo (2020) for 1990–2007, a difference of period and of margin: the estimates here cover 2005–2017, and the time-use analysis concerns the hours of employed fathers; the displacement estimates concern the employment-to-population ratio of all working-age adults. Dauth et al. (2021) find in German data that robots reduced hiring while incumbents kept their jobs and adjusted within the firm. Moreover, a commuting-zone long difference cannot separate a true increase in male labor-market attachment from compositional change (e.g., selective out-migration of non-employed men from exposed zones), so this coefficient cannot identify a "breadwinner-effort" response on the extensive margin. The estimate does establish what the paper's argument requires: the *sign* of the relative-wage shock (men's wages fall by more than women's), together with the *absence* of the male employment collapse that a pure labor-demand-displacement account would predict as the driver of the female withdrawal. The male-effort interpretation rests on the within-couple time-use evidence and the structural model; this aggregate employment coefficient carries none of it. The one quantity that does not sharpen under controls is the gender wage *gap* itself: it narrows significantly in the raw specification (-0.0019) but is imprecise once division fixed effects absorb regional trends, because both spouses' wages fall in level even as the husband's falls by more. The relative-wage channel and the male-effort response that the paper builds on are thus both present at Census scale.

Day-level margins. Table A17 decomposes the diary response into the probability of any market work on the diary day and the length of the working day conditional on working, the margins discussed in Section 3.3: the probability of any market work falls by 2.1 to 2.7 percentage points per robot (3 to 4 on weekdays) and the weekday working day shortens by 15 to 17 minutes, with men the mirror image at the ten-percent level.

Weekly hours worked. ATUS diaries capture a single day and record zero-work diaries for individuals sampled on a day off, which raises outcome noise and makes estimates sensitive to functional form in the presence of corner solutions. To verify that the findings are not an artifact of the one-day snapshot, Table 6 in the main text re-estimates the baseline using respondents' usual weekly hours at all jobs (IPUMS ATUS `UHRSWORKT`). For most respondents this is the figure reported at the final CPS interview, two to five months before the diary; it is re-asked at the ATUS interview only of respondents who changed jobs or employers in the interim or whose earnings were allocated. Because usual hours are collected for workers, these regressions condition on employment. The female estimates remain aligned with the diary-based results: a one-unit increase in robot exposure reduces women's usual working time by approximately 1.0 to 1.2 hours per week (Columns 4–6), consistent with the daily reduction of 16–19 minutes found in the diary estimates. The male coefficients, by contrast, are small and statistically insignificant in this measure (Columns 1–3), with point estimates near zero. This asymmetry has a natural reading: the increase in men's market effort that appears in the diary minutes is concentrated in same-day hours (overtime and additional shifts) which a "usual hours" figure, carried forward for most respondents from the CPS interview months earlier, largely misses. The usual-hours measure therefore corroborates the

Table A17: Day-level margins of the diary response: any market work on the diary day and the length of the working day (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
P(any market work on the diary day)						
Robots	0.010	0.013*	0.013	-0.023**	-0.027***	-0.021*
	(0.009)	(0.007)	(0.009)	(0.010)	(0.010)	(0.011)
Mean dep. var.	0.592	0.631	0.628	0.362	0.561	0.558
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Minutes of market work, working days						
Robots	10.037*	3.852	9.123	-8.490	-8.606	-7.422
	(5.904)	(3.901)	(5.900)	(6.679)	(6.389)	(6.655)
Mean dep. var.	453.5	459.4	455.4	377.0	386.2	382.5
Observations	2,321	4,097	2,304	2,518	2,753	2,466
P(any market work), weekdays						
Robots	0.016*	0.012	0.017*	-0.041***	-0.034***	-0.031**
	(0.010)	(0.009)	(0.010)	(0.013)	(0.012)	(0.013)
Mean dep. var.	0.850	0.903	0.901	0.531	0.824	0.823
Observations	1,926	3,171	1,792	3,524	2,371	2,123
Minutes on working days, weekdays						
Robots	11.329*	6.852	10.352	-15.674**	-17.270**	-15.471**
	(6.669)	(4.605)	(6.664)	(7.683)	(7.111)	(7.564)
Mean dep. var.	514.5	519.6	516.4	423.8	432.0	429.6
Observations	1,608	2,856	1,599	1,763	1,941	1,731

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering). The first block uses an indicator for any market work on the diary day; the second restricts to diary days with positive market work and uses the minutes worked; the third and fourth repeat both on Monday–Friday diaries. The diary minutes of Table 4 are the product of the two margins. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

female withdrawal but does not, on its own, detect the male intensive-margin response.

Weekend composition. Because the ATUS samples a single diary day and over-samples weekends (roughly half of all diaries fall on Saturday or Sunday), a concern is that the time-use estimates pick up weekend composition in place of a labor-supply response. Table A18 re-estimates the baseline on weekday (Monday–Friday) diaries only. The estimates come out *larger*: women's market-work withdrawal rises to 26–29 minutes per day (against 16–19 in the pooled sample) and men's increase to 13–18 minutes, with the childcare and leisure responses preserved. This is the pattern a labor-supply response implies, with the reallocation operating on working days, where market work is at stake, and it rules out a weekend-composition artifact.

Table A18: Robustness: weekday diaries only (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	17.904**	13.051**	18.263**	-26.549***	-28.755***	-26.396***
	(7.574)	(5.915)	(8.206)	(6.579)	(7.202)	(7.897)
Daily minutes childcare						
Robots	2.909	3.379	3.940	8.099**	14.253***	12.707***
	(2.669)	(2.685)	(2.638)	(4.085)	(4.899)	(4.645)
Daily minutes leisure						
Robots	-9.886**	-6.491	-14.077***	3.355	11.184***	9.453*
	(3.976)	(3.971)	(3.362)	(3.762)	(4.228)	(4.953)
Observations	1,926	3,171	1,792	3,524	2,371	2,123

Notes: Main 2SLS restricted to weekday (Monday–Friday) diaries, dropping the ATUS-oversampled weekend days. Effects are if anything larger than in the pooled sample, as a labor-supply response concentrated on workdays implies. Subsamples and FE as in Table 4; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Correction for sample selection bias. Selection into the sample is itself endogenous: automation shocks lower men's marriage-market value, reducing marriage and raising divorce risk (Anelli et al., 2024), so restricting the analysis to couples with young children could tilt the unobserved composition of the households that remain. To address this, I implement an Inverse Probability Weighting (IPW) procedure. I define the underlying at-risk population as all individuals in the ATUS aged 20–55 and estimate a probit model to predict the probability, $P(S_{ict} = 1)$, that an individual satisfies the baseline sample inclusion criteria. The selection equation is specified as:

$$P(S_{ict}=1 \mid \text{Robots}_{ct}^{IV}, \mathbf{X}_{ict}) = \Phi\left(\alpha + \gamma \text{Robots}_{ct}^{IV} + \mathbf{X}_{ict}'\beta + \delta_r\right), \tag{A2}$$

where $\mathbf{X}_{ict}$ represents a vector of demographic controls and δ_r denotes Census-region fixed effects. The selection equation includes the external instrument Robots_{ct}^{IV} (European robot adoption); the endogenous exposure measure never enters it. Using the instrument in the selection equation ensures that the correction isolates selection driven by exogenous global technological trends, uncontaminated by local unobservables that might simultaneously affect adoption and household stability. From the estimated parameters, I calculate the propensity score and construct inverse probability weights $w_{ict}^{IPW} = 1/\hat{p}_{ict}$. Table A19 presents the estimates corrected for selection. The market-work and leisure estimates are statistically indistinguishable from the baseline, so the reallocation does not reflect compositional shifts into and out of the couple sample.

Hourly-paid versus salaried workers. If the response reflects a labor-supply adjustment, it should be most visible among workers who can most readily change their hours, those paid by the hour. Merging the CPS-linked hourly-paid indicator onto the sample (Table A20), I interact exposure with hourly-paid status. The men's intensive-margin increase is significantly larger for hourly-paid workers (an additional 3 to 4 minutes per robot, significant in every column), as a flexible-hours adjustment implies. Women's withdrawal is present across pay structures and is slightly smaller among the hourly-paid, consistent with part of the female response operating on the extensive margin, scaling back or leaving a position, beyond marginal hour reductions. Either way, the response behaves like a reallocation of labor supply.

Concentration near the parity threshold. The norm interpretation makes a sharper prediction than an average withdrawal: the response should be concentrated among couples whose relative wages sit *near* the equal-earnings threshold, where the breadwinner constraint binds. I test this directly by interacting exposure with a *predetermined* near-parity indicator, equal to one when the wife's wage share predicted from spousal education and age gaps lies within 0.1 of parity, so that "near" is not the endogenous post-shock share (Table A21). The point estimates have the predicted sign: in every female column the near-parity interaction is negative, i.e. near-parity wives withdraw *more* (-2.7 to -4.9 minutes per robot on top of the baseline effect). But none of these interactions is statistically significant, so the test does not establish threshold concentration. The reason is a limitation of the cross-section, not evidence against the mechanism: predetermined education and age gaps predict the wage share only weakly, so the predicted share has little spread and fully 96% of couples fall in the "near" bin,

Table A19: Robustness to sample selection: effect of robots on market work, childcare, and leisure using inverse probability weighting (2SLS).

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	12.514**	5.782	12.973**	-16.053***	-19.490***	-17.808***
	(5.617)	(4.368)	(6.335)	(4.998)	(5.718)	(6.286)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.435	0.442	0.478	0.316	0.459	0.464
Daily minutes childcare						
Robots	-0.511	1.744	0.460	6.421**	9.876***	9.388***
	(2.125)	(1.969)	(2.013)	(2.754)	(3.127)	(3.251)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.260	0.194	0.278	0.258	0.248	0.263
Daily minutes leisure						
Robots	-3.448	-1.710	-8.130**	8.161**	13.376***	12.532***
	(4.557)	(3.478)	(3.451)	(3.429)	(3.505)	(4.239)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.321	0.271	0.332	0.226	0.319	0.329
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Sample includes married or cohabiting couples with at least one spouse working and children aged 10 or less. Columns 1 and 4 capture the extensive margin of the respondent (spouse is employed). Columns 2 and 5 capture the extensive margin of the spouse (respondent is employed). Columns 3 and 6 restrict to dual-earner households, capturing only the intensive margin of labor supply. Control variables include commuting zone and state-by-year FE, household demographics, day/month/holiday. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A20: Heterogeneity by hourly-paid status (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Hourly × Robots	3.738***	3.020**	3.564**	3.936**	3.597**	3.522**
	(1.424)	(1.269)	(1.401)	(1.540)	(1.479)	(1.683)
Robots	16.875***	11.481***	16.074***	-17.369***	-17.609***	-16.339**
	(5.849)	(3.931)	(5.699)	(6.047)	(6.035)	(6.514)
Daily minutes childcare						
Hourly × Robots	-0.581	-0.238	-0.628	-0.294	-0.516	-0.525
	(0.667)	(0.583)	(0.655)	(0.817)	(0.710)	(0.780)
Robots	1.458	2.916	1.878	7.519**	9.296***	10.232***
	(2.079)	(2.072)	(2.069)	(3.150)	(2.978)	(3.151)
Daily minutes leisure						
Hourly × Robots	-2.662**	-1.508	-2.434**	-0.836	-0.972	-1.123
	(1.047)	(0.941)	(1.077)	(1.001)	(0.924)	(1.053)
Robots	-7.565**	-3.583	-7.158**	10.787***	11.258***	10.690***
	(2.980)	(3.208)	(3.050)	(3.352)	(2.966)	(3.530)
Observations	3,429	5,821	3,360	4,169	4,398	3,942

Notes: 2SLS; robots and its interaction with an indicator that the respondent is paid by the hour (CPS-linked PAIDHOUR, IPUMS ATUS), both instrumented. Men's market-work increase is larger for hourly-paid workers; women's withdrawal is broad-based across pay structures. Subsamples and FE as in Table 4; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

leaving the interaction almost no contrast to exploit. The commuting-zone level, however, offers a predetermined moderator with real variation: the share of a zone's dual-earner couples sitting within 0.1 of parity, measured in the 2000 Census from 1.26 million couples, before the robot wave, and ranging from 0.39 at the 10th to 0.46 at the 90th percentile across zones. Interacting exposure with this (z-scored) near-parity density, with both terms instrumented (Table A22), yields directional but not decisive support on the headline margin: in the employed and dual-earner samples women's market-work withdrawal deepens by a further 4.3–6.6 minutes per robot per standard deviation of near-parity density, the predicted sign, but not statistically significant. Where the concentration *is* detectable is on the adjacent margins the mechanism also predicts: women's leisure expansion is significantly larger in near-parity-dense zones (+6.1 minutes per robot per standard deviation, dual-earner, 5%) and fathers' childcare falls significantly more (−4.1, 10%), the same reallocation pattern, concentrated where more couples sit near the threshold. The pattern is directional evidence for threshold concentration, though the work-margin interaction itself is insignificant; the per-couple threshold prediction remains untestable without panel data on couples' distance to parity. I therefore continue to treat the threshold reversals as descriptive cross-sectional patterns (Section 2.1) that motivate and restrict the structural model; I do not treat them as separately identified reduced-form results.

Table A21: Concentration of the effect near earnings parity (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Near × Robots	-3.469	-6.102**	-3.235	-4.872	-2.650	-4.425
	(3.594)	(2.379)	(3.129)	(4.496)	(4.817)	(5.098)
Robots	15.808**	14.920***	16.807***	-10.769	-15.797**	-11.825
	(6.280)	(4.528)	(6.487)	(6.574)	(7.248)	(7.744)
Daily minutes childcare						
Near × Robots	0.894	0.382	0.568	-1.995	-1.548	-0.396
	(1.341)	(1.300)	(1.323)	(3.094)	(2.319)	(2.180)
Robots	0.061	2.093	1.270	8.700**	11.271***	10.499***
	(2.334)	(2.360)	(2.280)	(4.039)	(3.617)	(3.516)
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS; robots and robots×Near both instrumented. Near = 1 if the couple's wage share predicted from *predetermined* spousal education and age gaps lies within 0.1 of parity (96% of couples fall in this bin, so the test has little contrast). The female near-parity interactions have the predicted (negative) sign but are not statistically significant. Controls and FE as in the main table; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A22: Concentration near parity: 2000-Census commuting-zone moderator (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
$\text{NearParity}_{2000} \times$ Robots	-0.105	-3.814	-0.100	0.646	-4.289	-6.587
	(5.074)	(4.718)	(4.928)	(3.809)	(4.181)	(4.543)
Robots	12.499	0.249	13.731	-14.543	-27.961**	-30.864***
	(13.017)	(12.654)	(12.660)	(10.231)	(10.935)	(11.781)
Daily minutes childcare						
$\text{NearParity}_{2000} \times$ Robots	-3.634	-0.613	-4.106*	0.047	-1.988	-4.148
	(2.241)	(2.270)	(2.201)	(3.034)	(3.260)	(3.922)
Robots	-7.271	1.064	-7.355	6.681	5.305	0.958
	(5.793)	(6.289)	(5.628)	(6.939)	(6.905)	(8.323)
Daily minutes leisure						
$\text{NearParity}_{2000} \times$ Robots	-4.660	0.216	-3.262	-0.071	3.951*	6.148**
	(3.375)	(2.857)	(3.150)	(2.305)	(2.157)	(2.444)
Robots	-15.913*	-3.278	-16.186*	6.051	19.702***	23.284***
	(8.708)	(7.650)	(8.258)	(6.148)	(5.605)	(6.314)
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS; robots and the interaction both instrumented by the EU5 shifter and its interaction. NearParity_{2000} is the z-scored share of the commuting zone's dual-earner couples whose wife's hourly-wage share lies within 0.1 of parity, measured in the 2000 Census from 1.26 million couples (predetermined w.r.t. the post-2000 robot wave; cross-zone spread 0.39 at the 10th to 0.46 at the 90th percentile). Unlike the couple-level proxy of Table A21, this moderator has real contrast. Controls and FE as in the main table; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The spouse's labor supply. The ATUS collects one diary per household but records, at the ATUS interview, the spouse's employment status and usual weekly hours (`SPEMPSTAT`, `SPUSUALHRS`). Table A23 uses them as outcomes in the baseline design: the spouse's employment where the spouse's extensive margin is free (columns 2 and 5) and the spouse's usual hours where the spouse is employed (columns 1, 3, 4 and 6). Neither moves. Husbands report no change in their wives' employment (+0.002, s.e. 0.010) and a small, insignificant decline in their usual hours (−0.35 and −0.37 hours per week, s.e. 0.29 and 0.26); wives report no change in their husbands' employment (+0.005, s.e. 0.007) or usual hours (−0.56 and −0.64, s.e. 0.51 and 0.57). Spouse-reported usual hours are recall reports and share the limitation of Table 6: they register the wife's withdrawal only weakly and the husband's increase not at all. Within-couple reallocation is therefore inferred from the gendered pattern of the respondents' diaries and from the structural model; it is never observed household by household.

Table A23: The spouse's labor supply as an outcome: employment and usual weekly hours (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Spouse employed (respondent employed)						
Robots		0.002			0.005	
		(0.010)			(0.007)	
Mean dep. var.		0.605			0.896	
Observations		6,593			4,891	
Spouse's usual weekly hours (spouse employed)						
Robots	-0.350		-0.367	-0.562		-0.644
	(0.286)		(0.260)	(0.511)		(0.568)
Mean dep. var.	35.8		35.6	45.0		44.4
Observations	3,980		3,707	6,901		4,245

Notes: 2SLS with the specification of Table 4 (same subsamples, controls, fixed effects, weights and commuting-zone–year clustering). Outcomes are the employment status and usual weekly hours of the respondent's spouse, reported by the respondent at the ATUS interview: in columns (1)–(3) the spouse is the wife, in columns (4)–(6) the husband. The spouse's employment is an outcome only where the spouse's extensive margin is free (columns 2 and 5); usual hours are defined for employed spouses (columns 1, 3, 4 and 6; hours-vary codes set to missing). Usual hours are recall reports, and the right benchmark is Table 6; diary minutes measure a different object. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Mechanism support: singles, childcare prices, education, and continuous norms. Four tables support Section 3.5 directly. Table A24 re-estimates the baseline for singles, whose market work and leisure do not respond and whose childcare falls where partnered mothers' rises. Table A25 reports the response of measured childcare prices, which move by about one percent per unit of exposure and in offsetting directions across care types. Table A26 interacts the shock with the wife's education, and Table A27 replaces the binary traditional-adherence split with the continuous, z-scored share of traditional adherents.

Table A24: Effect of robots on daily minutes of work, childcare, and leisure (2SLS). Sample: respondents with no spouse or unmarried partner.

	Single fathers	Single mothers
Daily minutes work		
Robots	-20.990	9.380
	(21.192)	(5.793)
R-squared	0.682	0.354
Daily minutes leisure		
Robots	12.273	-1.613
	(17.246)	(8.884)
R-squared	0.641	0.313
Daily minutes childcare		
Robots	-0.830	-8.304**
	(11.980)	(3.888)
R-squared	0.465	0.355
Observations	543	2,833

Notes: Controls include commuting zone and state-by-year FE, respondent's age/education, day, month, holiday. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A25: Childcare-cost channel: robot exposure on local childcare prices (2SLS)

Childcare price (log)	Coef. (Robots)	(SE)	t
Center-based, infant	+0.0091***	(0.0030)	3.04
Center-based, toddler	+0.0082**	(0.0037)	2.20
Center-based, preschool	+0.0101***	(0.0037)	2.69
Family care, infant	−0.0120***	(0.0046)	−2.60

Notes: 2SLS of the log county-level childcare price (DOL National Database of Childcare Prices, mapped to commuting zones, 2008–2018) on robot exposure, instrumented by the EU5 shifter, with commuting-zone and year fixed effects; SEs clustered at the CZ level (135 CZs, 1,228 CZ-year cells). Exposure moves prices only slightly and in offsetting directions, far too little to drive the female labor-supply response (see text).

Shock-based standard errors (AKM). Consistent with the identifying assumption of exogenous industry-level shifts, standard clustering at the commuting-zone level may understate standard errors if residuals are correlated across regions through shared sectoral

Table A26: Heterogeneity by the wife's education (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Wife coll. × Robots	-0.934	-0.957	0.030	1.240	1.520	2.331
	(1.882)	(1.540)	(1.855)	(1.147)	(1.610)	(1.581)
Robots	12.971**	9.288**	13.945**	-17.014***	-19.824***	-18.459***
	(5.448)	(3.755)	(5.806)	(4.630)	(5.524)	(6.061)
Daily minutes childcare						
Wife coll. × Robots	0.680	0.690	0.385	-2.287*	-1.948	-2.683**
	(0.697)	(0.562)	(0.690)	(1.167)	(1.316)	(1.351)
Robots	0.681	2.235	1.675	8.544***	11.371***	12.507***
	(2.133)	(1.957)	(2.061)	(2.715)	(3.069)	(3.320)
Daily minutes leisure						
Wife coll. × Robots	-1.324	-0.804	-1.176	-0.459	-0.345	-0.178
	(1.485)	(1.126)	(1.561)	(1.012)	(1.251)	(1.401)
Robots	-5.157	-3.503	-8.641***	6.600**	11.294***	9.907***
	(3.724)	(3.115)	(3.061)	(3.024)	(2.959)	(3.564)
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS; robots and its interaction with an indicator that the wife holds a tertiary degree, both instrumented. The market-work withdrawal does not differ by education (insignificant interaction), arguing against a pure robot–skill complementarity channel. Subsamples and FE as in Table 4; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A27: Heterogeneity by continuous (z-scored) traditional-adherence share (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
$\text{Trad}_z \times$Robots	-24.079	-12.759	-27.177	-35.623*	-29.698*	-36.589*
	(26.335)	(16.941)	(24.007)	(19.937)	(17.555)	(21.983)
Robots	1.756	3.126	1.595	-30.668***	-31.196***	-31.635***
	(13.608)	(9.039)	(12.840)	(9.848)	(9.833)	(11.494)
Observations	4,100	6,593	3,812	7,194	4,891	4,400

Notes: 2SLS; robots and its interaction with the continuous (z-scored) traditional-adherence share both instrumented. Subsamples and columns as in Table 7. Controls and FE as in the main table; SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

shocks. Households in geographically distant areas (such as Detroit, MI and Cleveland, OH) may experience correlated residuals because they are both exposed to the same global automotive industry shocks. I implement the exposure-robust inference procedure (AKM) to re-center inference on the industry-level shifts in place of the regional shares. Because the design has only a handful of effective shocks (the inverse-Herfindahl of the per-industry variance contributions is $g_{\text{eff}} \approx 2.8$–$5.5$ across specifications), I do *not* use Gaussian critical values, which are invalid with so few shocks (Adão et al., 2019), but instead t critical values with $g_{\text{eff}} - 1$ degrees of freedom, a conservative finite-shock correction. Table A28 reports the result. The female market-work withdrawal survives this stringent inference in the broad sample (Column 4, $t = -5.4$, significant at 5%) and conditioning on the respondent's employment (Column 5, also 5%), but *not* in the smallest, dual-earner cell (Column 6), where the corrected interval includes zero. Men's market-work increase survives at the 5–10% level, and the leisure responses, women's increase and men's decrease, survive in most columns. Women's childcare increase does *not* survive: its exposure-robust standard errors roughly double and it is insignificant throughout. The headline female withdrawal and the leisure reallocation are thus robust to exposure-robust inference even under the conservative finite-shock correction (outside the smallest cell), whereas the childcare margin is identified only under conventional clustering.

Clustering at the commuting-zone level. The baseline specifications cluster standard errors at the commuting-zone–year level, which allows arbitrary correlation within a zone–year cell but treats a zone's residuals as independent across years. Because robot exposure is highly persistent, serially correlated residuals within commuting zones could make that choice too liberal (Bertrand et al., 2004); the conservative alternative in a shares-based design is to cluster at the commuting-zone level, the level of the identifying shares (Goldsmith-Pinkham et al., 2020). Tables A29 and A30 re-report both headline tables with standard errors clustered by commuting zone (133–142 clusters, depending on the column); the samples and point estimates are identical by construction, so only the inference changes. The baseline results come out sharper: every female coefficient keeps its significance level (women's leisure in the broadest sample strengthens from the five- to the one-percent level), and the only cell that weakens is men's dual-earner market work, which slips from the five- to the ten-percent level ($p = 0.053$ against 0.027 under commuting-zone–year clustering). The norm-heterogeneity interactions are modestly noisier, as expected once serial correlation is absorbed: the women's market-work interaction remains significant among employed respondents ($p = 0.021$) and at the ten-percent level among dual earners ($p = 0.060$), but falls short of conventional significance in the spouse-employed sample ($p = 0.143$); the childcare

Table A28: Effect of robots on time use (2SLS). Adão et al. (2019) test for standard errors.

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	12.734**	8.967*	13.953*	-15.949**	-18.510*	-16.361
	(3.375)	(3.797)	(4.966)	(2.970)	(5.783)	(5.849)
g_{eff}	3.7	5.5	3.3	4.2	3.3	2.8
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.426	0.443	0.476	0.286	0.422	0.433
Daily minutes childcare						
Robots	0.854	2.466	1.771	6.579	9.686	10.092
	(2.182)	(1.338)	(2.179)	(5.801)	(7.075)	(7.213)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.229	0.180	0.253	0.239	0.231	0.246
Daily minutes leisure						
Robots	-5.494	-3.772*	-8.935**	6.205	10.996**	9.746*
	(2.783)	(1.376)	(2.487)	(2.497)	(2.863)	(3.279)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
R-squared	0.289	0.259	0.313	0.208	0.305	0.314
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Sample includes married or cohabiting couples with at least one spouse working and children aged 10 or less. Columns 1 and 4 capture the extensive margin of the respondent (spouse is employed). Columns 2 and 5 capture the extensive margin of the spouse (respondent is employed). Columns 3 and 6 restrict to dual-earner households, capturing only the intensive margin of labor supply. Control variables include commuting zone and state-by-year FE, household demographics, day/month/holiday. Point estimates are the main 2SLS coefficients; standard errors are the Adão et al. (2019) exposure-robust SEs, clustering shocks by industry×year. Because the effective number of shocks g_{eff} (inverse-Herfindahl of the per-industry variance contributions, reported in the first work block) is small, significance stars use t critical values with $g_{\text{eff}} - 1$ degrees of freedom in place of Gaussian critical values, a conservative finite-shock correction. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

interaction stays significant ($p = 0.036$), and the leisure interactions are unchanged or stronger. The amplification of the female withdrawal in traditional commuting zones therefore does not hinge on the clustering level, though its work-margin precision in the broadest sample does.

Table A29: Baseline estimates with standard errors clustered at the commuting-zone level.

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	12.734**	8.967**	13.953*	-15.949***	-18.510***	-16.361***
	(6.156)	(4.294)	(7.159)	(4.259)	(5.263)	(5.510)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Daily minutes childcare						
Robots	0.854	2.466	1.771	6.579**	9.686***	10.092***
	(2.488)	(1.881)	(2.150)	(3.022)	(2.906)	(2.880)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Daily minutes leisure						
Robots	-5.494	-3.772	-8.935***	6.205***	10.996***	9.746***
	(3.575)	(2.909)	(3.107)	(2.298)	(2.003)	(2.614)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: 2SLS. Specifications, samples, point estimates, and fit are identical to the main table (Table 4); only the clustering level changes. Standard errors in parentheses are clustered at the commuting-zone level (133–142 clusters, depending on the column), allowing arbitrary serial correlation of a zone's residuals across years (Bertrand et al., 2004). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

E Exposure by sex, industry, and occupation

This appendix reports in full the displacement evidence summarized in Section 3.5 of the main text: the sex-specific rebuild of the exposure measure and instrument, the split by the respondent's own industry, the couple-level predetermined potential exposures, and the occupational robot dose.

The alternative under test is that the female withdrawal reflects robots displacing women directly, so that no household response to the husband's labor market is involved. Commuting-zone exposure pools both sexes, so by itself it cannot say whose jobs the robots reach. Following Anelli et al. (2024), I first rebuild exposure and the EU5 instrument from the 1990 and 1970 Census microdata with industry employment shares computed separately for men and

Table A30: Norm heterogeneity with standard errors clustered at the commuting-zone level.

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Trad × Robots	-8.250	16.430	0.229	-26.982	-29.814**	-33.686*
	(23.236)	(14.589)	(20.094)	(18.432)	(12.867)	(17.910)
Robots	11.764*	11.360**	13.980**	-15.881***	-19.354***	-17.256***
	(7.096)	(4.511)	(7.079)	(4.053)	(5.110)	(5.374)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Daily minutes childcare						
Trad × Robots	1.123	0.998	-4.302	23.577**	9.266	7.705
	(6.290)	(9.094)	(7.013)	(11.238)	(6.531)	(7.821)
Robots	0.986	2.611	1.265	6.519**	9.949***	10.297***
	(2.626)	(2.548)	(2.333)	(2.776)	(2.942)	(2.900)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Daily minutes leisure						
Trad × Robots	-1.875	-27.504***	-5.131	5.166	15.066**	21.138**
	(15.867)	(8.729)	(11.532)	(7.341)	(7.589)	(10.274)
Robots	-5.714	-7.777**	-9.539**	6.192***	11.422***	10.308***
	(4.461)	(3.318)	(3.799)	(2.340)	(2.118)	(2.792)
Observations	4,100	6,593	3,812	7,194	4,891	4,400
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: 2SLS; robots and robots×Trad both instrumented (EU5 shifter and its interaction). Specifications, samples, and point estimates are identical to the norm-heterogeneity table (Table 7); only the clustering level changes. Standard errors in parentheses are clustered at the commuting-zone level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

for women (Table A32). Robotized employment is male: the automotive industry, which reaches 132 robots per thousand workers by 2019, was 21 percent female in 1990, basic metals 14 percent, while the industries where women predominate (textiles at 64 percent female, services at 53) are barely robotized (Table A31). Men's exposure alone reproduces every baseline estimate, with first-stage F statistics of 329–429 and similar magnitudes per standard deviation of exposure; women's exposure alone reproduces the female responses with a much weaker first stage (F of 37–55) and none of the male ones. The two measures cannot be entered jointly: because both sexes are exposed through the same handful of industries, they are correlated at 0.83 within the fixed effects, and the conditional first-stage F of the joint specification is at most six. A commuting zone's exposure is, to a first approximation, the exposure of its men.

Table A31: Who works in the robotized industries: 1990 female employment share and 2019 robot penetration

	Female share (1990)	Employment share (1990)	Robots per 1,000 workers (2019)
Automotive	21%	1.0%	132.0
Manufacturing, other	38%	1.4%	27.1
Electronics	39%	2.7%	18.9
Basic metals	14%	0.7%	16.4
Petrochemicals	30%	1.2%	11.4
Food and beverages	34%	1.3%	8.8
Metal products	23%	1.0%	6.5
Machinery	19%	1.5%	3.5
Non-metallic minerals	24%	0.5%	2.2
Other vehicles	21%	0.9%	1.1
Services	53%	72.7%	0.4
Furniture	23%	1.1%	0.3
Paper	39%	2.2%	0.2
Utilities	21%	1.3%	0.2
Textiles	64%	1.5%	0.2
Research	39%	0.4%	0.1
Mining	15%	0.6%	0.1
Agriculture and fishery	18%	1.8%	0.1
Construction	10%	6.2%	0.0

Notes: Industries in the IFR classification used throughout. Female share and employment shares from the 1990 Census 5% sample (employed, ages 16–64, all commuting zones); robots per 1,000 workers from the IFR operational stock in 2019 divided by 1990 industry employment, with unspecified robots allocated proportionally, as in the exposure measure.

Second, the ATUS records the respondent's industry, which allows a direct test of whether the response is concentrated among women who themselves work in robotized industries (Table A33). It is not concentrated there. Among employed women outside manufacturing, 94 percent of them, market work falls by 18 minutes per robot, childcare rises by 10 and leisure by 11; for women employed in manufacturing the withdrawal is four to five minutes smaller and the childcare and leisure increases about three minutes smaller, differences that

Table A32: Male versus female robot exposure (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Pooled exposure (Census-built)	23.855*	14.292	27.611*	-37.520***	-41.507***	-36.453**
	(13.565)	(10.026)	(14.677)	(11.657)	(13.981)	(15.222)
Men's exposure (alone)	18.405**	11.488*	20.774**	-23.965***	-25.759***	-22.543**
	(8.521)	(6.015)	(9.201)	(7.446)	(8.959)	(9.736)
Women's exposure (alone)	-28.010	-51.465	-20.671	-105.855**	-147.207***	-135.602**
	(75.069)	(68.989)	(78.881)	(53.940)	(56.901)	(64.097)
Observations	4,560	7,372	4,231	8,067	5,479	4,901
Daily minutes childcare						
Pooled exposure (Census-built)	0.842	6.123	3.335	18.261***	24.274***	26.748***
	(5.217)	(4.843)	(5.042)	(6.471)	(6.950)	(7.506)
Men's exposure (alone)	-0.404	3.158	1.327	10.783***	14.703***	16.156***
	(3.370)	(2.984)	(3.226)	(4.015)	(4.331)	(4.677)
Women's exposure (alone)	23.322	35.245	26.769	77.147**	85.991***	96.330***
	(24.681)	(25.834)	(24.153)	(31.197)	(27.983)	(31.514)
Observations	4,560	7,372	4,231	8,067	5,479	4,901
Daily minutes leisure						
Pooled exposure (Census-built)	-7.565	-5.810	-15.982**	14.342*	28.097***	23.989***
	(9.538)	(7.925)	(8.044)	(7.428)	(7.672)	(8.990)
Men's exposure (alone)	-6.003	-4.212	-11.353**	9.078*	18.015***	15.471***
	(5.962)	(4.886)	(4.942)	(4.752)	(4.812)	(5.671)
Women's exposure (alone)	35.769	15.235	6.620	52.521	86.945**	75.722*
	(49.788)	(46.260)	(49.649)	(33.163)	(38.560)	(43.817)
Observations	4,560	7,372	4,231	8,067	5,479	4,901

Notes: Each pair of rows is a separate 2SLS regression. All exposure measures in this table are rebuilt from the 1990 Census 5% microdata (employed, ages 16–64) in place of the County-Business-Patterns shares of the baseline, with the industry robot-penetration terms identical to the baseline; the first row pair (pooled shares) shows that this rebuild reproduces the baseline pattern. *Men's (women's) exposure* uses the commuting-zone industry employment shares of men (women) only, instrumented with the EU5 shifter built on the corresponding 1970 sex-specific shares. "Alone" enters one measure at a time. Controls, fixed effects, weights and commuting-zone–year clustering as in the main table. First-stage Kleibergen–Paap F: 308–388 (pooled), 329–429 (men's), 37–55 (women's). The two sex-specific measures are not separately identified: their correlation in the estimation sample is 0.87 (0.83 after the fixed effects) and, entered jointly (two endogenous regressors, two instruments), the Sanderson–Windmeijer conditional first-stage F is at most 6. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

are significant at the one-percent level for childcare. The response is thus carried by women whose own jobs robots do not reach, as a household reallocation to the husband's labor-market shock implies; direct displacement of women would predict the opposite concentration. Men's increase in market work is likewise not confined to men in manufacturing (the interactions are small and insignificant), consistent with an adjustment that runs through the local labor market; displacement at the plant would concentrate where the robots are. One caveat applies: industry is measured after the shock, so a woman displaced from manufacturing into services is counted outside manufacturing; with six percent of employed women in manufacturing, however, such moves cannot account for a withdrawal of this size across the other 94 percent.

Table A33: Is the response concentrated among respondents who themselves work in robotized industries? (2SLS)

	Males		Females	
	Resp. emp.	Dual earner	Resp. emp.	Dual earner
Daily minutes work				
Robots × respondent in manufacturing	1.841	0.472	4.848*	3.798
	(1.191)	(1.660)	(2.513)	(2.474)
Robots	8.028**	13.688**	-18.322***	-16.368***
	(3.841)	(5.816)	(5.410)	(5.872)
Observations	6,593	3,812	4,891	4,400
Daily minutes childcare				
Robots × respondent in manufacturing	-0.501	-0.685	-3.042***	-2.775***
	(0.506)	(0.705)	(0.997)	(1.063)
Robots	2.722	2.155	9.569***	10.097***
	(1.994)	(2.173)	(2.655)	(2.853)
Observations	6,593	3,812	4,891	4,400
Daily minutes leisure				
Robots × respondent in manufacturing	-0.629	-0.701	-2.699*	-2.986**
	(1.175)	(1.453)	(1.473)	(1.438)
Robots	-3.451	-8.542***	10.891***	9.752***
	(3.060)	(3.076)	(2.931)	(3.501)
Observations	6,593	3,812	4,891	4,400

Notes: 2SLS on the employed columns of the main table (the respondent's industry is observed for the employed): robots and its interaction with an indicator that the respondent's main job is in manufacturing (Census industry codes 1070–3990), both instrumented with the EU5 shifter and its interaction. Manufacturing employs 15% of the employed men and 6% of the employed women in the sample. The uninteracted coefficient is the response of respondents outside manufacturing. Controls, fixed effects, weights and commuting-zone–year clustering as in the main table. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Third, the same conclusion follows when exposure is assigned to each *spouse* instead of to each place. Because the ATUS does not record the spouse's occupation or industry, and a displaced worker's current job is itself an outcome, I assign both spouses the *predetermined* exposure of their demographic cell, the potential-wage logic applied to exposure: the 1990 industry mix of the person's sex–education–age–state cell, directly and through the cell's occupations, times the year's IFR penetration, instrumented with its EU5 counterpart (Table A34). Neither spouse's cell exposure produces the female withdrawal: the husband's cell exposure leaves women's market work unchanged, and men in exposed cells *raise* their market work, the same direction as the local estimate. The respondent's own current occupation tells the same story at the job level: among employed women, a higher robot dose of one's own occupation comes with *more* market work (Table A35). The female withdrawal is therefore neither direct displacement (it is absent exactly where robots reach women's own industries, occupations and demographic cells) nor a response to the husband's individual exposure; it is a response to the *local* labor market moving against the male-breadwinner occupation mix, the reading the couple-level first stage of Section 3.2 imposes and the commuting-zone estimates identify.

F Industrial automation versus trade shocks

This section reports the comparison of the robot shock with import competition from China summarized in Section 3.4 of the main text.

Import competition from China depressed manufacturing employment and wages for men across U.S. labor markets (Autor et al., 2013) and, like the robot shock, lowered the marriage-market value of young men, with marriage and fertility falling in exposed areas (Autor et al., 2019). Men are also disproportionately concentrated in occupations exposed to tradable industries (Figure A4), so they bear the initial burden of both disruptions.

The two shocks nonetheless reach the household budget through different margins, and for trade the predicted household response is ambiguous. Trade arrives as industry-level contraction, with plant closures and permanent layoffs on the extensive margin (Halla et al., 2020): the loss of male income activates a consumption-smoothing, added-worker motive that pushes wives toward market work (Besedeš et al., 2021), while the degradation of the husband's economic status might activate the same identity-based withdrawal observed under automation, as the household attempts to preserve traditional gender roles in the face of

[2] The tradable share is computed as the pooled, person-weighted probability for each occupation o that a worker i is employed in a tradable industry:

$$\hat{p}_o = \frac{\sum_{i \in o} w_i \, \mathbf{1}\{\text{Industry}_i \in T\}}{\sum_{i \in o} w_i},$$

where T includes Agriculture/Forestry/Fishing, Mining, Manufacturing, Transport/Communications/Utilities, and Wholesale Trade.

Table A34: Whose exposure? Market work on the husband's and the wife's predetermined potential exposure (2SLS)

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Cell industry mix						
Husband's potential exposure (alone)	6.904**	5.105**	8.573***	2.114	-0.833	-1.562
	(3.301)	(2.322)	(2.881)	(3.188)	(3.807)	(4.033)
	[F 1002]	[F 1651]	[F 811]	[F 1590]	[F 1479]	[F 1480]
Wife's potential exposure (alone)	21.680*	9.939	25.913**	8.035	12.417	8.118
	(11.554)	(9.331)	(10.181)	(9.454)	(10.829)	(11.920)
	[F 235]	[F 410]	[F 210]	[F 768]	[F 664]	[F 514]
Husband's, joint	5.453*	4.538**	6.921**	1.445	-1.884	-2.331
	(3.175)	(2.277)	(2.873)	(3.243)	(3.685)	(3.957)
Wife's, joint	17.431	6.252	19.554**	6.619	13.881	9.969
	(11.022)	(9.166)	(9.670)	(9.478)	(10.139)	(11.356)
Observations	4,090	6,584	3,807	7,181	4,880	4,393
B. Cell occupational mix (dose through jobs)						
Husband's potential exposure (alone)	-8.584	-10.321	2.159	5.493	-2.939	-7.825
	(15.779)	(11.536)	(14.676)	(9.782)	(11.555)	(12.196)
	[F 1277]	[F 1350]	[F 1092]	[F 1713]	[F 1435]	[F 1443]
Wife's potential exposure (alone)	30.219	6.600	47.160*	7.073	18.814	13.486
	(26.545)	(16.662)	(24.835)	(15.291)	(20.730)	(22.893)
	[F 568]	[F 936]	[F 445]	[F 719]	[F 510]	[F 382]
Husband's, joint	-20.255	-15.636	-12.925	3.519	-12.025	-16.215
	(15.806)	(11.812)	(14.925)	(11.406)	(13.264)	(13.958)
Wife's, joint	41.084	14.642	54.114**	4.939	27.555	25.667
	(26.605)	(17.155)	(25.019)	(17.620)	(23.390)	(25.715)
Observations	4,090	6,584	3,807	7,181	4,880	4,393

Notes: 2SLS with the controls and fixed effects of Table 4 on the baseline estimation sample. Each spouse's potential exposure is the 1990 employment mix of their sex–education–age–state cell over the 19 IFR industries (Panel A: the cell's direct industry mix; Panel B: the mix reached through the cell's occupations), multiplied by the IFR U.S. robot penetration of the year; the instrument replaces U.S. penetration with the EU5 average, as in the baseline. Cells are predetermined, so neither displacement-induced industry switching nor the spouse's unobserved current job enters the measure. In brackets, the first-stage F of the own instrument; the joint rows instrument both exposures with both shifters. Standard errors clustered by the husband's education–age–state cell. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A35: The occupational robot dose: market work on the dose of the respondent's own occupation (2SLS)

	Males		Females	
	Employed	Dual earner	Employed	Dual earner
Own-occupation dose (employed respondents)				
Own-occupation robot dose	0.504	-0.349	4.481**	3.636
	(1.417)	(1.714)	(2.068)	(3.097)
	[F 164]	[F 176]	[F 120]	[F 138]
Observations	6,447	3,711	4,821	4,342
Own dose and the spouse's predetermined cell dose, joint				
Own-occupation robot dose	0.509	-0.512	4.534**	3.722
	(1.336)	(1.778)	(2.142)	(2.704)
Spouse's cell dose	7.039	52.467**	-2.939	-6.572
	(15.904)	(22.652)	(10.653)	(11.084)
Observations	6,442	3,709	4,810	4,335

Notes: 2SLS with the controls and fixed effects of Table 4, employed respondents of the baseline estimation sample. An occupation's dose is its national 1990 industry mix over the 19 IFR industries times the IFR U.S. robot penetration of the year (instrument: the EU5 average penetration); the respondent's current occupation (mapped to occ1990dd with Dorn's crosswalks) carries it, so the estimates are read descriptively; the current occupation is post-shock. The joint panel adds the spouse's predetermined cell dose (Table A34, Panel B). Standard errors clustered by the respondent's occupation (upper panel) and by the husband's education–age–state cell (joint panel). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

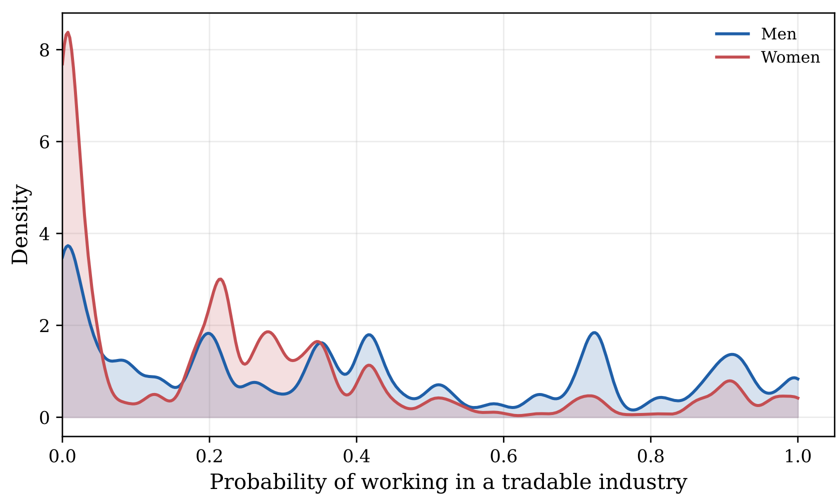

Figure A4: Kernel densities, by sex, of the occupation-level probability of working in a tradable industry, $\hat{p}_o$ (defined in the footnote). Workers aged 18–64, 1990 five-percent Census sample (Ruggles et al., 2024), person-weighted; the person-weighted means are 0.39 for men and 0.24 for women.

manufacturing decline.

Industrial robots are instead a task-based shock that substitutes for labor mainly at the intensive margin, devaluing specific manual tasks while often leaving the firm operationally viable (Acemoglu and Restrepo, 2020). Automation therefore depresses male potential wages and hours without the immediate unemployment of a plant closure, so the insurance motive should weigh less against the identity motive. Comparing the two shocks tests whether the identity mechanism generalizes to trade: estimates directionally similar to the robot ones would mean that the identity wedge dampens or overrides the added-worker effect even during broad manufacturing contraction.

Two exercises separate the shocks. First, I re-estimate the baseline replacing robot exposure with commuting-zone exposure to Chinese imports, built following Autor et al. (2013), so the two shocks, one task-based and one trade-based, can be compared on the same within-household outcomes. Second, I re-estimate the robot coefficients controlling for the trade shock directly, so that the automation response cannot be a stand-in for the broader manufacturing decline that accompanied import competition.

Measurement of the trade shock. The trade measure draws on several sources. Detailed bilateral trade data by product and year are retrieved from the UN Comtrade Database. Domestic production and absorption values used to normalize the shocks are sourced from the U.S. Census Bureau and the NBER-CES Manufacturing Industry Database. Historical industry-specific employment at the commuting-zone level is derived from the Decennial Census/ACS microdata and County Business Patterns (CBP), following the regional definitions and industry concordances established by Autor et al. (2013).

U.S. imports from China expanded rapidly from the early 1990s (driven by China's internal reforms, rural-to-urban migration, and improved access to foreign technology) and accelerated after WTO accession in 2001 granted permanent normal trade relations. As shown in Figure A5, imports from China rose through the 1990s, entered a phase of rapid growth following WTO accession, and stabilized around 2014.[3]

Following Autor et al. (2013), I extract annual bilateral export flows from China by product to the United States and to a set of other high-income destinations. Product-level values are mapped to U.S. industries using standard NAICS concordances and aggregated to construct industry-level Chinese import exposure. Local exposure to import competition is a shift-share in the cumulative growth of Chinese import penetration, weighted by the commuting zone's historical industrial composition in 1990. Specifically, industry-level imports from China are

[3]The ATUS sample used to investigate the effect of import competition is restricted to the period 2003–2014 to align with the peak years of the trade disruption.

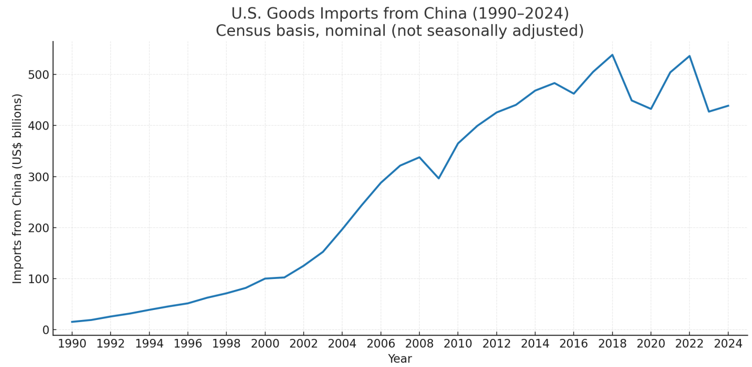

Figure A5: U.S. goods imports from China, 1990–2024 (Census basis, nominal).

mapped to 397 four-digit SIC manufacturing industries. The local exposure is defined as:

$$\text{Trade}_{ct} = \sum_j \frac{\text{Emp}_{jc,1990}}{\text{Emp}_{c,1990}} \times \frac{\Delta\text{Import}_{jt}^{Ch\to US}}{\text{Absorption}_{j,1990}}, \tag{A3}$$

where $\Delta\text{Import}_{jt}^{Ch\to US} = \text{Import}_{jt}^{Ch\to US} - \text{Import}_{j,1991}^{Ch\to US}$ represents the growth in U.S. imports from China in industry j between the base year and year t. The denominator is the sector's 1990 domestic absorption, defined as total production plus imports minus exports. This normalization expresses the shock as a change in the penetration ratio, so absolute volume drops out. I rely on the change in exposure relative to the baseline to isolate the structural displacement affecting labor demand, a strategy analogous to the robot exposure measure defined in Equation 1.[4]

To address the potential endogeneity of U.S. import penetration (the concern that import growth might reflect U.S. domestic demand shocks instead of Chinese supply factors), I instrument U.S. exposure using the industry-level growth of Chinese exports to eight other high-income economies: Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland. The use of these alternative markets isolates the supply-driven component of China's export growth, such as productivity gains and falling trade costs, that is orthogonal to U.S. economic conditions. The instrument Trade_{ct}^{IV} replaces the U.S. import term with these foreign import flows and utilizes 1980 employment shares as weights. The use of 1980 shares further ensures that the measure is not contaminated by anticipation effects or mechanical correlations with the labor market adjustments occurring in the 1990s and 2000s.

[4]Unlike the trade measure, the robot exposure variable does not require explicit differencing from a base year because the operational stock of industrial robots in the early 1990s was negligible across most industries; thus, the contemporaneous stock is effectively equivalent to the cumulative adoption since the beginning of the period.

Gender-oriented decomposition. Trade exposure varies across sectors with different gender intensities, so it can be decomposed into shocks that load on male-dominated and on female-dominated industries, which sharpens the comparison with the male-biased robot shock. Using the historical female share of employment in each industry-commuting zone pair, $f_{cj,1990}$, and its complement, $1 - f_{cj,1990}$, I decompose the overall exposure into gender-specific shocks:

$$\text{Trade}^g_{ct} = \sum_j \omega^g_{cj,1990} \times \frac{\text{Emp}_{jc,1990}}{\text{Emp}_{c,1990}} \times \frac{\Delta\text{Import}^{Ch\to US}_{jt}}{\text{Absorption}_{j,1990}}, \quad g \in \{f, m\}, \tag{A4}$$

where $\omega^f_{cj,1990} = f_{cj,1990}$ and $\omega^m_{cj,1990} = 1 - f_{cj,1990}$. This decomposition identifies two distinct sources of variation: Trade^m_{ct} targets sectors such as heavy manufacturing and basic metals where men prevail, while Trade^f_{ct} loads on sectors like textiles and electronics with higher female employment shares.

Table A36 reports summary statistics for these aggregate and gender-specific trade measures. The aggregate exposure to Chinese import competition (Trade_{ct}) has a mean of 0.006 and a standard deviation of 0.010. Given the construction of the variable as a weighted average of industry-level penetration ratios, a one-standard-deviation increase corresponds to approximately a 1 percentage point rise in the share of Chinese imports in domestic absorption. Decomposing this exposure reveals the gendered nature of the disruption. The male-specific component (Trade^m_{ct}), which isolates exposure in male-dominated industries, averages 0.003. The female-specific component (Trade^f_{ct}) is slightly smaller, with a mean of 0.002.

Table A36: Summary statistics of trade exposure measures.

Variable	Mean	Std. Dev.	Min.	Max.	N
Trade_{ct}	0.006	0.010	0	0.163	168,414
Trade^m_{ct}	0.003	0.005	0	0.059	168,414
Trade^f_{ct}	0.002	0.004	0	0.044	168,414

China shock results. Table A37 provides the initial set of comparative results using the aggregate Chinese import penetration measure. In contrast to the divergent, identity-consistent response generated by robot exposure, rising import competition depresses male market hours sharply while the female response is muted. The 2SLS estimates for the two-earner sample (Column 3 for men, Column 6 for women) indicate that a 1 percentage point increase in import penetration reduces male market work by approximately 58 minutes per day, alongside a smaller, statistically imprecise reduction in female market work of around

22 minutes.

The female response to the aggregate trade shock, however, departs from the automation pattern. Unlike the robot shock, the aggregate China shock does not produce a robust increase in women's childcare: the dual-earner point estimate is positive but statistically imprecise (Column 6), and in the broader sample women's market work moves up while their non-market time falls (Column 4), a configuration closer to the standard added-worker response. This contrast reinforces the view that trade operates largely as a generalized negative wealth shock, driven by the extensive-margin loss of manufacturing jobs; the relative-wage, identity-consistent channel that characterizes task-based automation is largely absent.

The decomposition into Trade_m and Trade_f in Table A38 looks at the identity-norm mechanism more closely. The male-specific shock (Trade_m) is a mirror-image benchmark for the robot analysis: it disproportionately targets male-dominated manufacturing hubs and thereby raises the wife's relative potential wage share. Under Trade_m, the dynamics mirror the automation results: men in two-earner households experience a significant reduction in their leisure time (between 118 and 153 minutes per 1 percentage point increase), which is consistent with an effort to maintain their absolute household contribution as their sector declines. On the other hand, the results for Trade_f, which loads on sectors where women prevail, show a reversal. In this case, the shock to the wife's relative earnings is accompanied by a large and significant withdrawal of the husband from market work (approximately 244 minutes per 1 percentage point increase) and an increase in his leisure time. These decomposed per-percentage-point coefficients should be read as local-linear slopes; evaluating them at a full percentage point would overstate them: the gender-specific trade shocks have standard deviations well below one percentage point (Table A36), so the relevant within-support change is a fraction of these figures. Scaled to a one-standard-deviation move they remain large but no longer exceed the daily time budget; I therefore emphasize the *sign* and the contrast with the robot shock, and I put no weight on the literal per-point magnitudes.

Joint estimation of robot and trade shocks. As a final robustness check, I estimate a joint specification that includes both robot exposure and Chinese import competition simultaneously. This strategy isolates the impact of task-based automation while holding constant the broader effects of manufacturing decline and globalization. Table A40 presents these results.

For market work and leisure, the qualitative pattern on robot exposure is preserved when the trade shock is controlled for. The robot coefficients themselves are, however, larger in this specification than in the headline table (e.g., male two-earner market work of 35.53 minutes against 13.95 in the baseline), and it would be a mistake to read that gap as evidence

Table A37: Effect of the China shock on daily minutes of work, childcare, and leisure (2SLS).

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Trade	-6,756***	-6,755***	-5,823***	2,377**	-1,475	-2,180
	(1,789)	(2,571)	(1,502)	(1,189)	(1,360)	(1,431)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.405	0.418	0.454	0.246	0.387	0.400
Daily minutes childcare						
Trade	1,295	1,679	907.8	-1,327**	488.8	1,042
	(881.7)	(1,495)	(845.0)	(658.2)	(780.0)	(797.7)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.212	0.167	0.225	0.224	0.217	0.224
Daily minutes leisure						
Trade	-441.6	635.1	-551.5	-2,004**	-495.3	-593.8
	(1,222)	(1,215)	(1,133)	(862.5)	(937.4)	(984.8)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.284	0.258	0.307	0.195	0.280	0.290
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Trade coefficients are expressed per unit of the import-penetration share (a 0–1 scale); a one-percentage-point increase equals 0.01 units, so the per-percentage-point effects quoted in the text are the tabulated coefficients $\times 0.01$ (e.g. $-5{,}823 \times 0.01 \approx -58$ minutes). Sample: married or cohabiting couples with at least one spouse working and children aged 10 or less. Controls include state-by-year FE, demographic/household controls, day/month/holiday FE. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A38: Effect of the decomposed China shock on daily minutes of work, childcare, and leisure (2SLS).

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Trade_{Male}	4,766	-3,365	6,638	7,546	3,211	-866.4
	(8,137)	(7,261)	(7,019)	(5,538)	(7,103)	(7,532)
Trade_{Fem}	-23,909**	-11,932	-24,429**	-5,990	-9,100	-4,246
	(12,098)	(11,144)	(11,083)	(8,465)	(11,263)	(12,057)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.406	0.418	0.455	0.246	0.387	0.400
Daily minutes childcare						
Trade_{Male}	7,675**	2,100	6,029*	-6,030*	-4,445	-6,444
	(3,863)	(2,827)	(3,351)	(3,197)	(3,453)	(3,985)
Trade_{Fem}	-8,339	1,103	-6,851	6,179	8,405	12,977**
	(5,512)	(4,656)	(4,952)	(4,987)	(5,426)	(6,118)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.211	0.167	0.224	0.223	0.217	0.223
Daily minutes leisure						
Trade_{Male}	-11,753**	-3,408	-15,303***	-38.3	4,903	3,676
	(5,724)	(4,026)	(5,223)	(3,859)	(4,300)	(4,769)
Trade_{Fem}	16,420*	6,792	21,510**	-5,038	-9,149	-7,351
	(8,924)	(6,558)	(8,382)	(6,068)	(6,953)	(7,584)
Observations	3,716	6,244	3,466	6,814	4,608	4,093
R-squared	0.286	0.258	0.310	0.195	0.279	0.289
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Sample includes married or cohabiting couples with at least one spouse working and children aged 10 or less. Columns 1 and 4 capture the extensive margin of the respondent (spouse is employed). Columns 2 and 5 capture the extensive margin of the spouse (respondent is employed). Columns 3 and 6 restrict to dual-earner households, capturing only the intensive margin of labor supply. Trade coefficients are per unit of the import-penetration share (a 0–1 scale; multiply by 0.01 for the per-percentage-point effect). All specifications use 2SLS. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

that "trade and automation exert opposing pressures." Almost all of the difference is a *sample* effect: the robot–trade joint estimation is necessarily restricted to the 2003–2014 import-penetration window, and the robot coefficient estimated on that same window with *no* trade control is already 32.96 (Table A39). Adding the trade shock moves it only to 35.53. The trade control thus accounts for about one-eighth of the change; the rest is the 2003–2014 composition of the sample. The same window restriction explains why the female robot coefficient, while stable in magnitude (Column 4, −16.07; Column 6, −15.72), loses precision relative to the full-sample estimates; spatial correlation between the two shocks plays no role, since the robot-only window estimates in Table A39 are already imprecise before any trade control is added. The substantive point that survives is the one the joint estimation is meant to make: conditioning on the China shock leaves the robot coefficients essentially unchanged in sign and magnitude, so automation's effect on household time use is not a by-product of concurrent trade exposure. The shift in leisure, a reduction for men and an increase for women, is likewise preserved.

However, the results for childcare time are less robust in the joint estimation. While the baseline specification indicated a shift toward female childcare, the coefficients in Table A40 are smaller in magnitude and statistically insignificant for both spouses. This sensitivity indicates that the specific shift of non-market time toward childcare is difficult to disentangle from the broader disruptions associated with trade shocks, or that the high spatial correlation between the two shocks increases the standard errors. Nevertheless, the stability of the results for market work and leisure supports the finding that automation shifts intra-household labor supply toward a traditional gendered allocation independently of concurrent trade patterns.

Beyond the diverging margins of adjustment, the relative noise in the trade estimates likely reflects the conceptual distinction between aggregate regional disruption and localized productivity shifts. As documented by Autor et al. (2013) and Dauth et al. (2021), trade shocks frequently trigger general equilibrium spillovers (local housing market declines, a collapse in non-tradable service demand) that exert a simultaneous negative wealth effect on both spouses. These spillovers add aggregate noise on top of the within-household margins of interest. Industrial robots, by comparison, act as a scale-expanding productivity change within surviving plants, and they rarely eliminate whole local industries (Adachi et al., 2024), so automation isolates the shift in the within-couple relative price of time more cleanly. The dollar-value flows behind trade exposure are also more sensitive to exchange-rate movements and reporting lags than counts of installed robot units, which contributes to the trade estimates' lower precision.

Table A39: Robots on market work: full sample vs. China-window vs. joint with trade (2SLS)

	Full 2003–19	China window 2003–14	+ trade (joint)
Men Spouse emp.	12.73**	26.05***	28.82***
Men Resp emp.	8.97**	16.85**	18.53**
Men Dual Earner	13.95**	32.96***	35.53***
Women Spouse emp.	-15.95***	-15.21*	-16.07*
Women Resp emp.	-18.51***	-10.26	-10.31
Women Dual Earner	-16.36***	-16.09	-15.72

Notes: Each cell is the 2SLS robot-exposure coefficient on daily market-work minutes (EU5-instrumented). Column 1 is the baseline full sample (Table 4); column 2 restricts to the 2003–2014 China-shock window with robots as the only endogenous regressor; column 3 adds the China shock (the joint horse race, Table A40). The change from column 1 to column 3 comes overwhelmingly from the window restriction (col. 1→2); the trade control adds little (col. 2→3). SEs clustered at the CZ–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

G Estimation details

This appendix collects the inputs and the procedural detail behind the estimator of Section 4.2: the simulation inputs, the free-knot test of the parity restriction, the efficient-weight re-estimation and over-identification test, the equal-weights re-estimation, the simulation and optimization protocol, the profiled identification of the fixed costs, the sensitivity of the estimates to the moments, and the untargeted extensive margin.

Simulation inputs. Table A41 reports the ATUS summary statistics that parameterize the log-normal wage and fertility distributions of Section 4.2.

The location of the kink. Fixing the knot of (11) at parity is a restriction the data can be asked about, and I test it directly. Re-estimating the two-segment model with a *free* knot $k \in [0.30, 0.70]$ on the same micro sample, block by block, locates the kink where the norm predicts (Figure A6): the maternal-childcare V, the sharpest reversal, puts it at $\hat{k} = 0.52$ with a respondent-bootstrap 95% confidence interval of $[0.44, 0.58]$, tightly bracketing the equal-earnings point, and female market work puts the point estimate at $\hat{k} = 0.53$, with 77% of bootstrap draws within 0.05 of parity, though its full interval is wide; the two gently sloped male profiles do not pin the location, as expected. Imposing $k = 0.5$ is nearly costless in fit: it raises the residual sum of squares by at most 0.03% relative to the best free knot in any block, so the parity restriction never binds against the data. Applied to the simulated benchmark, the same free-knot fit checks the model itself: the wedge $\phi(s)$ is smooth, with no built-in threshold at 0.5, yet the simulated profiles place their turning points in the near-parity region ($\hat{k}$ between 0.41 and 0.51, at or to the left of the threshold, as a smooth convex wedge implies).

Table A40: Joint effects of robots and import competition on daily minutes of work, childcare, and leisure (2SLS).

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Daily minutes work						
Robots	28.821***	18.526**	35.527***	-16.068*	-10.311	-15.721
	(7.791)	(7.957)	(7.562)	(8.644)	(12.610)	(14.097)
Trade	-7,301***	-7,257**	-6,425***	3,991*	180.8	-776.0
	(2,532)	(3,616)	(2,035)	(2,096)	(2,600)	(2,717)
Observations	3,014	4,993	2,809	5,388	3,660	3,274
R-squared	0.402	0.416	0.453	0.260	0.392	0.409
Daily minutes childcare						
Robots	-0.187	1.565	-0.233	-1.714	2.861	2.424
	(3.216)	(3.896)	(3.106)	(5.111)	(6.305)	(7.057)
Trade	-933.0	-176.8	-1,428	-1,599	-608.1	104.6
	(1,238)	(1,750)	(1,162)	(1,057)	(1,419)	(1,471)
Observations	3,014	4,993	2,809	5,388	3,660	3,274
R-squared	0.213	0.181	0.233	0.231	0.221	0.228
Daily minutes leisure						
Robots	-5.077	-1.144	-10.582**	9.669	12.295*	16.820**
	(6.568)	(4.922)	(5.298)	(6.508)	(6.921)	(7.400)
Trade	-872.3	104.7	-627.3	-2,034	483.9	499.7
	(1,763)	(1,723)	(1,473)	(1,519)	(1,680)	(1,722)
Observations	3,014	4,993	2,809	5,388	3,660	3,274
R-squared	0.272	0.252	0.293	0.195	0.288	0.295
Spouse employed	✓		✓	✓		✓
Respondent employed		✓	✓		✓	✓

Notes: Sample includes married or cohabiting couples with at least one spouse working and children aged 10 or less. Columns 1 and 4 capture the extensive margin of the respondent (spouse is employed). Columns 2 and 5 capture the extensive margin of the spouse (respondent is employed). Columns 3 and 6 restrict to dual-earner households, capturing only the intensive margin of labor supply. The Trade coefficients are per unit of the import-penetration share (a 0–1 scale; multiply by 0.01 for the per-percentage-point effect); the robot coefficients are per robot per 1,000 workers. Control variables in all specifications include state-by-year fixed effects, demographic/household controls, day/month/holiday FE. Standard errors clustered at the commuting-zone–year level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A41: Summary statistics for simulation inputs

Variable	Mean	Std. Dev.
Female wages (w_f)	16.49	9.60
Male wages (w_m)	20.07	11.17
Number of children (n)	2.14	1.00

The reversal at parity is thus a location that the data select and that the model generates without a built-in threshold.

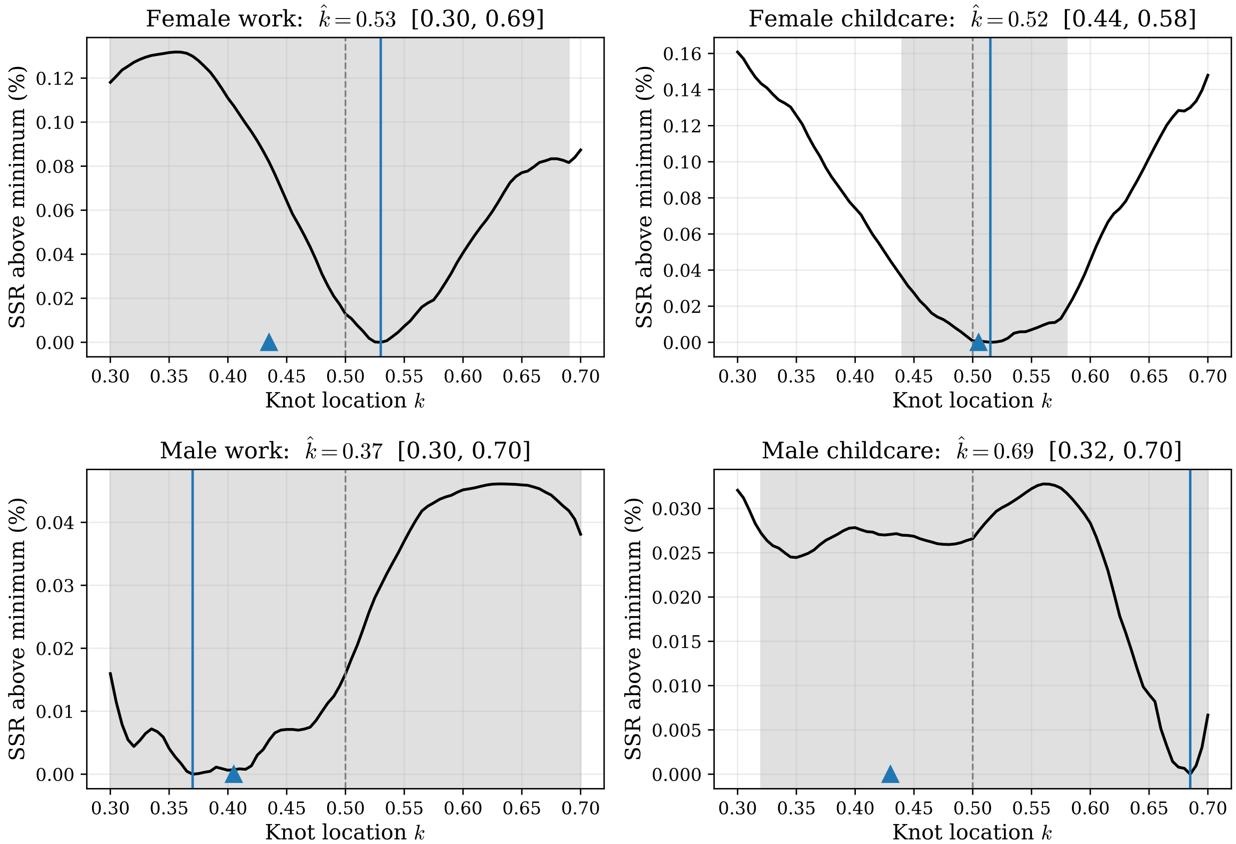

Figure A6: Freely estimated kink location of the two-segment time-use profiles. Each panel profiles the residual sum of squares of the continuous two-segment fit (in percent above its free-knot minimum) as the knot k moves over $[0.30, 0.70]$, estimated on the structural micro sample (26,621 respondents, $0.25 \le s < 0.75$). The solid blue line marks the SSR-minimizing knot $\hat{k}$, the grey band its respondent-bootstrap 95% confidence interval ($B = 1{,}000$), the dashed grey line parity ($k = 0.5$), and the blue triangle the knot estimated on the benchmark model's simulated profiles, whose smooth wedge $\phi(s)$ has no built-in threshold. Maternal childcare locates the kink tightly at parity ($\hat{k} = 0.52$, CI $[0.44, 0.58]$); female work's point estimate sits at parity ($\hat{k} = 0.53$, 77% of bootstrap draws within 0.05 of parity) with a wide interval; the gently sloped male profiles do not pin the location. Imposing $k = 0.5$ raises the SSR by less than 0.03% in every block.

Efficient weighting and the over-identification test. Because this weighting is non-optimal by design, a Hansen over-identification test (which presumes the efficient weight matrix) is a secondary check here. I nonetheless implement it: I estimate the full 12×12 covariance $\hat{\Omega}$ of the data moment vector by a household bootstrap of the ATUS sample (26,621 respondents, $B = 3{,}000$) and re-estimate under the efficient weight $W^* = \hat{\Omega}^{-1}$ (Table A42). The estimated covariance confirms the conditioning problem (condition number $\approx 2{,}400$), and the efficient weight, loading on the precisely-estimated *level* moments and down-weighting the high-variance tail *slopes*, sacrifices the gentle curvature the design relies on: the female-work inverse-V survives, but the childcare curvature shrinks toward zero (the maternal-care V to a hundredth of its empirical amplitude, the paternal inverse-V to a faint one). This is the concern that motivated the self-normalizing weight: diagonal weighting preserves the identifying curvature that the data moments themselves carry. The over-identification test also rejects at the efficient weight ($J = 46.2$, χ^2_2, $p < 0.001$); a test of this kind rejects virtually any parsimonious static structural model at this sample size, and here the rejection is driven by the tail moments the efficient weight both down-weights and fails to reproduce. The model is over-identified (twelve moments for ten parameters, one of them disciplined externally).

Table A42: Efficient (optimal) weighting and the Hansen over-identification test

Block	Data slopes (β^L, β^R)	Benchmark (self-normalizing)	Efficient weight $\hat{\Omega}^{-1}$
Female work	$(+0.22, -0.13)$ inv-V	$(+0.15, -0.23)$ inv-V	$(+0.10, -0.09)$ inv-V
Female childcare	$(-0.15, +0.20)$ V	$(-0.03, +0.06)$ V	$(-0.003, +0.001)$ near-flat
Male work	$(+0.01, +0.21)$ rise	$(-0.09, +0.13)$ V	$(-0.07, +0.06)$ V
Male childcare	$(+0.01, -0.04)$ inv-V	$(+0.02, -0.03)$ inv-V	$(+0.008, -0.006)$ faint inv-V
Implied $\mathbb{E}[\theta]$	—	0.43	0.52
Hansen J (χ^2_2)	—	—	46.2 ($p < 0.001$)

Notes: Two-segment slopes of the simulated profiles under the benchmark self-normalizing weight and under the efficient weight $W^* = \hat{\Omega}^{-1}$, where $\hat{\Omega}$ is the full 12×12 covariance of the data moment vector (four levels and eight kinked slopes) estimated by a household bootstrap of the ATUS sample (26,621 respondents, $B = 3{,}000$; condition number of $\hat{\Omega} \approx 2{,}400$, confirming the ill-conditioning that motivates the self-normalizing weight). The efficient weight is estimated with the bargaining weight left free (no cap), so the Hansen J is a proper efficient-GMM statistic with $12 - 10 = 2$ degrees of freedom (all ten parameters are estimated). The efficient weight loads on the precisely-measured levels and down-weights the high-variance tail slopes: the female-work inverse-V survives, but the childcare curvature is shrunk toward zero (the maternal-care V to slopes of -0.003 and $+0.001$ against -0.15 and $+0.20$ in the data, the paternal inverse-V to a faint $+0.008/-0.006$), and the implied female Pareto weight drifts above parity. The over-identification test rejects, as it does for essentially any parsimonious static structural model at this sample size, driven by precisely the tail curvature the efficient weight discards. This is why the paper adopts the self-normalizing weight: it preserves the identifying curvature instead of manufacturing it.

Equal weights. Table A43 re-estimates the benchmark under equal weights $(w_\ell, w_\beta, \omega) = (1, 1, 1)$. The headline shapes, including the male-childcare inverse-V, survive, so they do not depend on the chosen weighting (Section 4.2).

Simulation and optimization. Moments are simulated on $N = 100{,}000$ households drawn as described above, with the bargaining-weight shocks z_i drawn once and held fixed across all parameter evaluations (common random numbers), so that $\hat{m}(\Theta)$ is a smooth, deterministic function of Θ and simulation noise is negligible relative to the data moments. The objective

Table A43: Robustness of the structural shapes to EQUAL objective weights

Block	Data slopes (β^L, β^R)	Benchmark $(6, 10, 2)$	Equal weights $(1, 1, 1)$
Female work	$(+0.22, -0.13)$ inv-V	$(+0.15, -0.23)$ inv-V	$(+0.15, -0.31)$ inv-V
Female childcare	$(-0.15, +0.20)$ V	$(-0.03, +0.06)$ V	$(-0.04, +0.10)$ V
Male work	$(+0.01, +0.21)$ rise	$(-0.09, +0.13)$ V	$(-0.09, +0.16)$ V
Male childcare	$(+0.01, -0.04)$ inv-V	$(+0.02, -0.03)$ inv-V	$(+0.02, -0.03)$ inv-V
Level+shape SSR (full N)	—	5.6	5.6

Notes: Two-segment slopes of the simulated profiles under the benchmark objective weights $(w_\ell, w_\beta, \omega_{\text{M-care}}) = (6, 10, 2)$ and under equal weights $(1, 1, 1)$ (re-estimated identically otherwise). The three reversals, including the demanding male-childcare inverse-V, are reproduced under equal weights, so they do not depend on the chosen weighting; under both weightings the model gives male market work a V where the data are flat below parity and rising above it.

(13) is non-convex; it is minimized in two stages, a global differential-evolution search followed by a local derivative-free polish (BOBYQA; Powell, 2009), restarted from several seeds, with the winning candidate re-evaluated at the full simulation sample. Standard errors are obtained by a household block bootstrap that re-draws the ATUS sample, recomputes the data moments, and re-optimizes the model from the point estimate by a local (derivative-free) polish. Because this re-optimization is local (no full global re-search runs at each draw), the resulting standard errors are best read as local-curvature measures; the profiled identification analysis provides the global counterpart and, for the wedge parameters, confirms a sharply pinned optimum.

Profiled identification of the fixed costs. The profiling of Section 4.2 extends to the fixed care costs. Re-optimized fits between $\tilde{t}_f = 0.04$ and the estimate lie within 0.3 SSR units of each other with all four shapes intact, while pushing the wife's fixed cost to zero costs more than one SSR unit and loses the paternal-care reversal. The level of the fixed costs is thus loosely pinned; the shapes require only that the wife's fixed cost not be small. The symmetric positivity floor on both fixed costs, a regularity condition that any present parent provides some baseline custodial care, is retained in the estimation bounds and never binds.

Sensitivity to the moments. I document the moment-to-parameter mapping with the sensitivity matrix of Andrews et al. (2017), $\Lambda = -(G'WG)^{-1}G'W$, computed at the estimate with the estimator's own weight matrix and standardized by the bootstrapped standard deviation of each moment (Table A44). The matrix confirms the mapping described in Section 4.2: every free parameter draws at least 97% of its absolute sensitivity from the *slope* moments, with the levels contributing almost nothing; both wedge parameters load most heavily on the arms of male market work (above all the right-of-parity slope), with the right-of-parity arm of female work and the maternal-childcare slopes next, the profiles the wedge bends, while the remaining slopes pin down the preference and bargaining-heterogeneity parameters. The wedge estimates move with the curvature; the levels barely register.

Table A44: Sensitivity of the wedge parameters to the estimation moments (Andrews et al., 2017)

Moment	Wedge scale χ	Wedge curvature κ
Female work: level	$+0.01$	$+0.01$
Female work: slope β^L / β^R	-0.36 / -0.68	-0.07 / -0.16
Female childcare: level	$+0.00$	-0.00
Female childcare: slope β^L / β^R	$+0.39$ / -0.21	-0.00 / $+0.09$
Male work: level	$+0.00$	-0.00
Male work: slope β^L / β^R	-1.31 / $\mathbf{-1.76}$	$+0.03$ / $\mathbf{-0.64}$
Male childcare: level	-0.00	-0.00
Male childcare: slope β^L / β^R	-0.03 / $+0.04$	$+0.00$ / -0.11
Share of $\lvert\Lambda\rvert$ on slope moments	99.7%	99.3%

Notes: Standardized Andrews–Gentzkow–Shapiro sensitivities $\Lambda = -(G'WG)^{-1}G'W$, computed at the benchmark estimate with the estimator's own (diagonal, self-normalizing) weight matrix; each entry is the movement of the parameter, in its own units, per one-standard-deviation perturbation of the data moment (moment standard deviations from the household bootstrap of Table A42). All ten parameters are estimated; the full 10×12 matrix is in `SMM/sensitivity_results.json`. Bold marks each parameter's largest-magnitude entry. Both wedge parameters are moved almost exclusively by the *slope* moments, each most by the arms of male market work, the profile the wedge bends through the wife's withdrawal, with levels contributing under 2%: the curvature identifies the wedge.

The untargeted extensive margin. Figure A7 plots the extensive-margin validation of Section 4.3: the participation profile across the wife's wage share under both data definitions, against the benchmark and the two no-norm alternatives.

H Policy counterfactuals

Section 4.5 prices four policy instruments on the estimated household. This appendix records how each instrument enters the solved model, the validation run before any number was produced, the incidence conventions and the bracket they imply, the dose–response paths behind the headline doses, and the bootstrap uncertainty. The implementation lives in `SMM/policy_lib.py`, the checks in `SMM/policy_validate.py`, and the results in `SMM/policy_counterfactuals.json` and `SMM/policy_ci.json`.

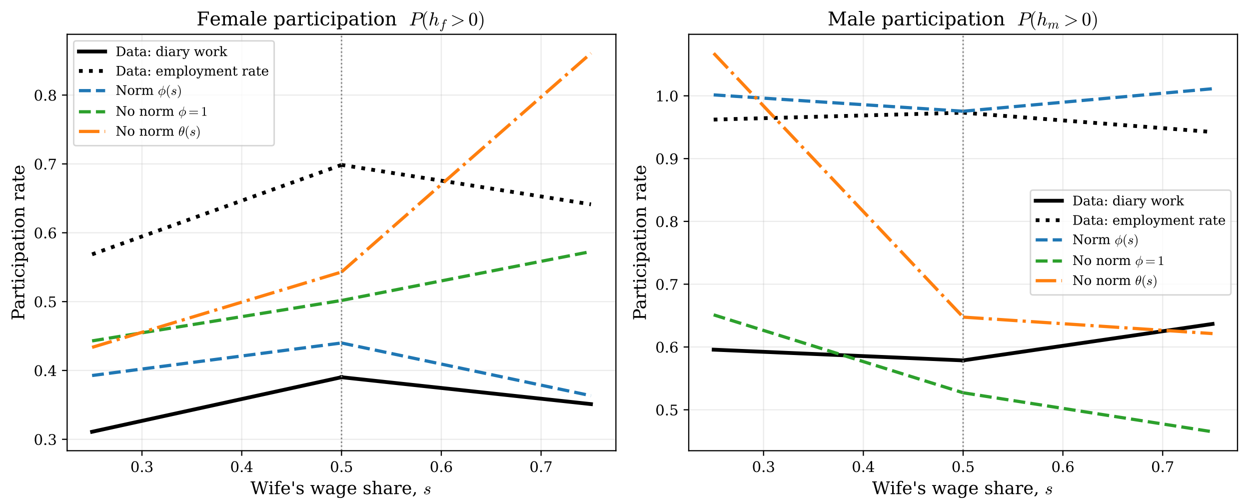

Figure A7: Untargeted extensive margin: participation rate $P(h_i > 0)$ by the wife's wage share, as two-segment fits kinked at parity. The two black curves are the data under the two definitions, diary-day work (solid) and the labor-force employment rate (dotted), and both trace the female inverse-V. Only the benchmark norm (blue) reproduces this profile; the wage-rising weight (orange) and the wedge-off re-estimate (green) are both monotone, with female participation rising in s, male participation falling, and no reversal at parity, so neither alternative bends the extensive margin where the data do. Participation is not used in estimation.

Implementation. Every experiment consumes the benchmark estimate and the solved twelve-regime model of Section J; nothing is re-estimated. The wedge parameter enters the solver only through the wife's effective wage $\phi(s)\,w_f$, so a per-household array $\chi_i = -\ln(\phi_i^{\text{eff}})/s_i^{\kappa}$ implements any effective-wage schedule $\phi_i^{\text{eff}}\,w_f$, with the closed forms, the KKT validity conditions, and the welfare-argmax regime selection untouched; a constant array reproduces the scalar benchmark to machine precision. Wage insurance and childcare provision need no such construction, since the husband's wage and the fixed care costs are ordinary solver inputs. Lump-sum instruments, child allowances among them, are outside this set: household consumption has no intercept in the twelve regime solutions, every first-order condition changes if one is added, and under logarithmic utility such transfers move the allocation through income effects alone. The experiments are therefore confined to instruments that scale prices and time costs, which the three spending policies and the phase-out all do.

Validation. Three checks precede every reported number. First, at dose zero each policy arm reproduces the plain post-shock solve exactly (largest allocation difference 4×10^{-16}, identical no-regime households). Second, the household-level identities (time budgets, budget positivity, non-negativity of hours and care) hold with zero violations in every arm at every dose, and the no-regime share never exceeds its baseline of 0.15 percent. Third, on random subsamples of roughly 250 households per arm, and on every household whose effective wage the gross-incidence phase-out drives to the floor while the wife keeps working, the solver's allocation is verified against direct numerical maximization of the reduced first-stage program; across some 3,400 households the largest shortfall is 4×10^{-15} welfare units. The floored-wage workers are the one configuration that looks anomalous and is analytically right: with a Pareto-weight draw near zero, the wife's leisure carries almost no weight in the household objective, and work at any positive wage is optimal.

Incidence conventions. A credit paid on the wife's gross earnings gives her $(\phi(s) + \sigma)\, w_f$ per hour; a credit that the display friction consumes at the same rate as her market earnings gives her $(1 + \sigma)\, \phi(s)\, w_f$. The main text uses the gross convention throughout and the choice matters little for the credit: the price of an hour under the norm is \$7.40 in the first convention and \$4.20 in the second, both far below the efficient household's \$162. The phase-out is where the conventions diverge, because τ subtracted from $\phi(s)$ can exhaust the effective wage while τ applied to the net wage cannot: under the gross convention $\tau = 0.21$ puts 83 percent of exposed wives past a 100 percent effective marginal rate, participation falls from 0.39 to 0.16, and 14.7 weekly hours are destroyed at \$1.60 of revenue each, while under the net convention the loss is 2.4 hours at \$4 each. The efficient household loses 1.1 hours at \$124 each under either convention, so the norm's amplification of the excess burden is a factor between 30 and 78. Every sign pattern holds in all 60 bootstrap re-estimates under both conventions.

Wage insurance in detail. Figure A8 traces the dose–response. The insured wage path retraces the main text's shock experiment in reverse, so full replacement returns the exposed couples to the pre-shock allocation by construction; the policy content is the ledger and the split of the response across regimes. At the Reemployment Trade Adjustment Assistance dose $r = \frac{1}{2}$ the norm household adds 0.3 weekly hours of female work and the efficient household sheds 0.5; at $r = 1$ the figures are $+0.5$ and -0.9, the outlay is \$68 per couple per week (18.3 percent of the restored wage bill), and the share of exposed couples above parity falls from 0.84 to its pre-shock 0.75. Insurance is the one instrument whose female-hours yield exists only because of the wedge, and its price per hour, \$134 under the norm, reflects the inframarginal payments to every displaced husband.

Childcare provision in detail. Figure A9 traces coverage from zero to the full custodial load under the proportional incidence of the main text (thick lines) and a mother-only variant (thin lines). At $\rho = \frac{1}{2}$ the freed 14.3 weekly hours split, under the norm, into 29 percent

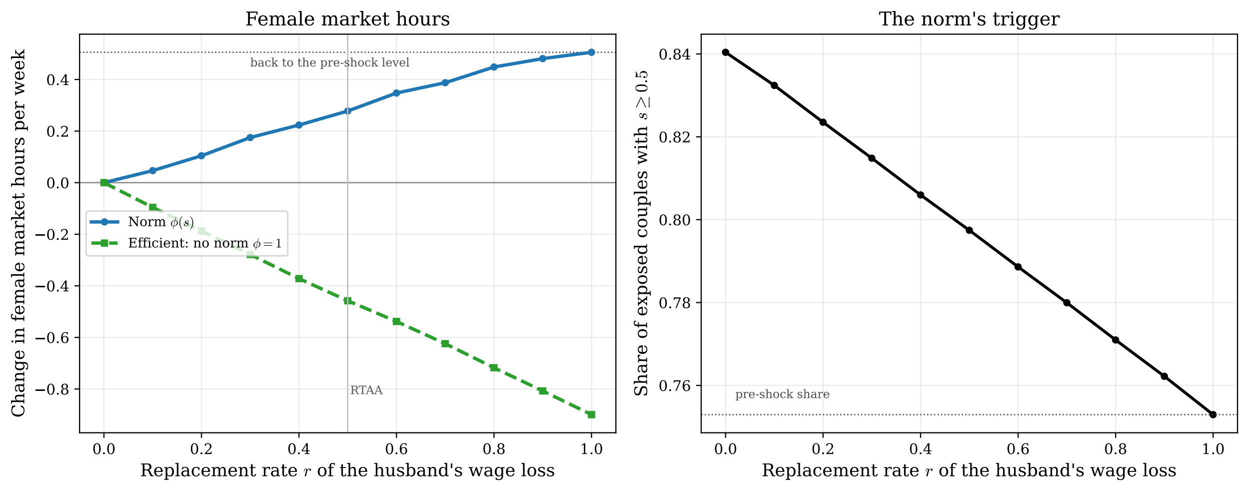

Figure A8: Wage insurance for displaced husbands: replacement rate r of the shocked bottom-tercile wage loss. Left: change in the exposed wives' market hours from $r = 0$, norm (blue) and efficient (green) regimes on a common axis; the dotted line marks the change that restores the pre-shock level, reached exactly at full replacement. Right: the share of exposed couples with $s \geq 0.5$, the norm's trigger, which is a property of wages alone and identical across regimes; full replacement returns it to the pre-shock share. The vertical line marks the RTAA design $r = \frac{1}{2}$.

market work, 17 percent quality care, and 54 percent leisure; the efficient split is 52, 8, and 40 percent. Mother-only slots of the same coverage free 9.8 hours, all hers, raise her market work by 4.7 weekly hours at \$104 of outlay, and close the care gap entirely near $\rho = 0.9$, since the fixed component of the gap is the part provision removes. Repairing the shock's damage to her hours takes coverage of 0.06 (proportional, \$19 per week) or 0.05 (mother-only, \$11). Provision is priced at the childcare workers' median wage in 2000, \$7.20 per hour (median usual weekly earnings of \$288 for full-time childcare workers in the CPS, at forty hours), with one caregiver per child; group care at staff ratio g divides every outlay by g and leaves the hours unchanged.

Bootstrap uncertainty. Table A45 pushes each of the 60 household-bootstrap re-estimates through every policy at its headline dose, under that draw's own parameter vector: the draw's wedge defines the effective-wage arithmetic, its fixed care costs define the coverage, and its pre-shock solve defines the restoration target. The headline sign patterns, the credit's advantage under the norm, the insurance sign flip, the damped daycare yield, and the amplified excess burden of the phase-out, each hold in all 60 draws.

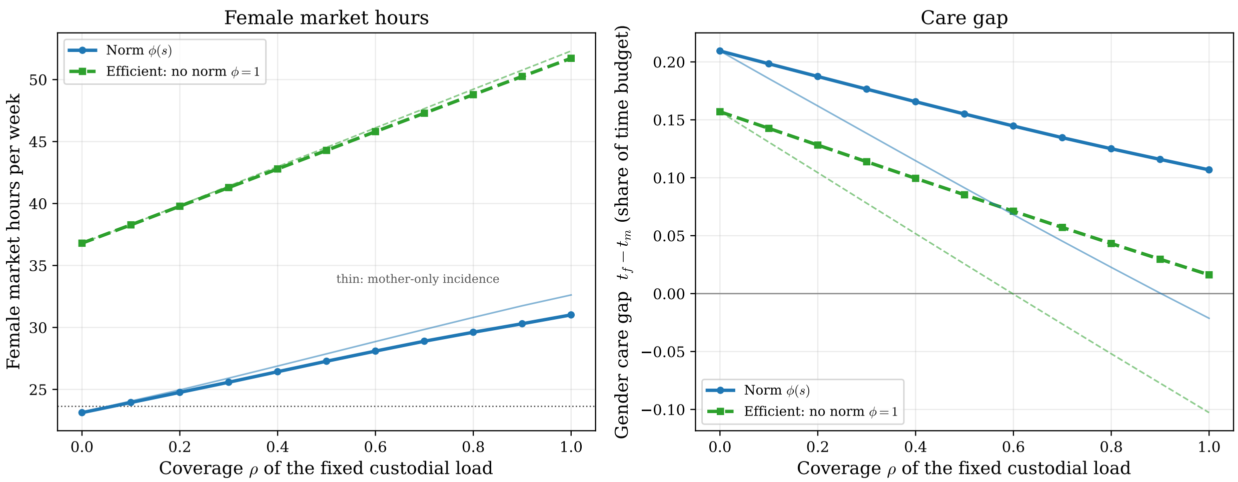

Figure A9: Public childcare as in-kind provision covering a share ρ of the fixed custodial time costs $\tilde{t}_i$, in the post-shock world. Left: the exposed wives' market hours; right: the gender care gap. Thick lines reduce both parents' fixed loads proportionally, thin lines the mother's alone; blue is the norm regime and green the efficient regime, at the same estimated parameters. The dotted line is the pre-shock level of female hours. Provision raises female hours and narrows the care gap in both regimes; the norm regime routes about half as much of the freed time into market work.

Table A45: Policy counterfactuals: bootstrap 90-percent intervals ($B = 60$)

	Δ female hours/week		Dollars per hour	
Instrument (dose)	Norm	Efficient	Norm	Efficient
Credit, gross ($\sigma = 0.10$)	$[+5.7, +7.8]$	$[+0.3, +0.5]$	$[6.5, 9.2]$	$[147, 202]$
Credit, net ($\sigma = 0.10$)	$[+0.9, +1.4]$	$[+0.3, +0.5]$	$[4.0, 5.3]$	$[147, 202]$
Wage insurance ($r = \frac{1}{2}$)	$[+0.2, +0.5]$	$[-0.5, -0.4]$	$[64, 147]$	$[42, 69]$
Wage insurance ($r = 1$)	$[+0.4, +0.9]$	$[-1.0, -0.7]$	$[71, 158]$	$[45, 74]$
Childcare provision ($\rho = \frac{1}{2}$)	$[+3.7, +4.6]$	$[+7.0, +8.1]$	$[33, 41]$	$[19, 22]$
Phase-out, net ($\tau = 0.21$)	$[-3.0, -2.0]$	$[-1.2, -0.8]$	$[3.2, 4.5]$	$[116, 162]$
Phase-out, gross ($\tau = 0.21$)	$[-16.8, -12.6]$	$[-1.2, -0.8]$	$[1.1, 2.7]$	$[116, 162]$

Notes: 5th and 95th percentiles across the 60 household-bootstrap re-estimates of Table 8, each pushed through the policy at its headline dose with the wedge on (norm) and off (efficient). Dollars per hour are absolute fiscal flows per weekly hour of female market work moved; the direction of the movement is the sign of the corresponding hours column. The wage-insurance interval for the efficient regime prices hours removed. Every headline sign pattern obtains in all 60 draws.

I Proofs of the propositions

I.1 Statement and proof of Proposition A1

A1 (Primitives). Preferences are defined as in (8) with parameters $\delta_i > 0$, $\mu_i > 0$, $\gamma_i > 0$, $\alpha \in (0,1)$, wages $w_i > 0$, number of children $n \in \mathbb{N}$, and fixed time costs $\tilde{t}_i \geq 0$ such that the time budget is feasible ($1 - n\tilde{t}_i > 0$) for $i \in \{f, m\}$.
A2 (Norm wedge). The function $\phi : [0,1] \to (0,1]$ is C^2, strictly decreasing, and log-convex in s (e.g., $\phi(s) = \exp[-\chi s^\kappa]$ with $\chi > 0, \kappa > 1$). Because s is defined on exogenous potential wage shares ($s = w_f/(w_f + w_m)$), $\phi(s)$ acts as a constant parameter during the household's time-allocation optimization. Therefore, household consumable income c is strictly linear in (h_f, h_m), ensuring the mapping $(h_f, h_m) \mapsto \log(c)$ is strictly concave on the feasible set.
A3 (Feasible sets). For any (h_f, h_m) such that $0 \leq h_i \leq 1 - n\tilde{t}_i$, each player's second-stage choice set for t_i is nonempty, convex, and compact. The joint feasible set for (h_f, h_m, t_f, t_m) is also nonempty, convex, and compact.

Proposition A1 (Existence and uniqueness)**.** *Under Assumptions A1–A3:*

(i) ***Stage 2 (Childcare).*** *For any pair of market hours (h_f, h_m), the second-stage game in (t_f, t_m) is a diagonally strictly concave game. Therefore, it admits a unique Nash equilibrium (t_f^*, t_m^*), and the equilibrium correspondence $(t_f^*, t_m^*)(h_f, h_m)$ is single-valued and continuous.*

(ii) ***Stage 1 (Market work).*** *Let $V(h_f, h_m) = \theta\, v_f\big(h_f, h_m, t^*\big) + (1-\theta)\, v_m\big(h_f, h_m, t^*\big)$ denote the household objective induced by the unique Stage 2 equilibrium. Under A2, V is strictly concave on the feasible set of (h_f, h_m); therefore, the Stage 1 problem has a unique maximizer (h_f^*, h_m^*).*

Hence, the two-stage semi-cooperative problem has a unique subgame-perfect equilibrium $(h_f^*, h_m^*, t_f^*, t_m^*)$.

Stage 2. Fix (h_f, h_m). Player i solves a strictly concave problem in t_i:

$$v_i(t_i; t_{-i}, h_f, h_m) = \log(c) + \mu_i \log\big(1 - h_i - (t_i + \tilde{t}_i)n\big) + \gamma_i \log\big((1+t_f)^\alpha (1+t_m)^{1-\alpha} n\big),$$

over a nonempty, convex, compact interval for t_i (A3). The second derivative in own strategy,

$$\frac{\partial^2 v_i}{\partial t_i^2} = -\frac{\mu_i n^2}{\big(1 - h_i - (t_i + \tilde{t}_i)n\big)^2} - \frac{\gamma_i}{(1+t_i)^2} \times \begin{cases} \alpha & \text{if } i = f, \\ 1-\alpha & \text{if } i = m, \end{cases}$$

is strictly negative, so each payoff is strictly concave in own action. The Jacobian of first-order conditions has a symmetric part whose diagonal entries dominate the off-diagonal terms (the cross–partial from $\log q$), implying diagonal strict concavity (DSC). By Rosen's theorem (Rosen, 1965), a unique Nash equilibrium (t_f^*, t_m^*) exists; by the implicit function theorem and DSC, (t_f^*, t_m^*) is continuous in (h_f, h_m).

Stage 1. Define $V(h_f, h_m) = \theta v_f\big(h_f, h_m, t^*(h)\big) + (1-\theta) v_m\big(h_f, h_m, t^*(h)\big)$. The feasible set for (h_f, h_m) is convex and compact (A3). The only non-affine term in (h_f, h_m) is $\log(c)$ with $c = \phi(s) w_f h_f + w_m h_m$. Because s is the exogenous potential wage share, $\phi(s)$ is independent of (h_f, h_m). Therefore, c is a linear combination of market hours, and the mapping $(h_f, h_m) \mapsto \log(c)$ is strictly concave on the feasible set. All remaining terms in V are concave in (h_f, h_m), and composition with the single-valued continuous correspondence $t^*(h)$ preserves concavity. Hence V is strictly concave and attains a unique maximizer (h_f^*, h_m^*). Combining with the unique (t_f^*, t_m^*) from Stage 2 yields a unique subgame-perfect equilibrium. □

I.2 The fully-cooperative benchmark and the invariance of the childcare ratio

The model comparison of Section 4.4 includes a fourth, most-demanding no-norm competitor: a *fully-cooperative* household that chooses all four time uses jointly, so that the Pareto weight θ enters the childcare margin directly, which the semi-cooperative timing rules out. This appendix records the closed-form solution of that benchmark and proves the invariance result that explains why even a fully-cooperative $\theta(s)$, given every opportunity to bend the childcare profiles, cannot reproduce the paternal-care reversal.

The fully-cooperative household solves

$$\max_{h_f, h_m, t_f, t_m \geq 0} \theta\, u_f + (1-\theta)\, u_m, \qquad u_i = \delta_i \log c + \mu_i \log l_i + \gamma_i \log\big((1+t_f)^\alpha (1+t_m)^{1-\alpha} n\big),$$

with $c = \phi(s) w_f h_f + w_m h_m$ and $l_i = 1 - h_i - (t_i + \tilde{t}_i)n$, i.e. the same primitives as the benchmark but with a single joint objective in place of the two-stage timing. Under the normalization $\delta_f = \delta_m = 1$, write $G(\theta) \equiv \theta\gamma_f + (1-\theta)\gamma_m$. In the interior regime, where both spouses work ($h_i > 0$) and both provide quality care ($t_i > 0$), the first-order conditions are

$$\underbrace{\frac{\alpha\, G(\theta)}{1+t_f} = \frac{\theta\mu_f\, n}{l_f}}_{\text{care, } f}, \qquad \underbrace{\frac{(1-\alpha)\, G(\theta)}{1+t_m} = \frac{(1-\theta)\mu_m\, n}{l_m}}_{\text{care, } m}, \qquad \underbrace{l_f = \frac{\theta\mu_f\, c}{\phi(s) w_f}}_{\text{hours, } f}, \qquad \underbrace{l_m = \frac{(1-\theta)\mu_m\, c}{w_m}}_{\text{hours, } m}. \tag{A5}$$

The care first-order conditions are the object the semi-cooperative timing suppresses: here the *social* marginal benefit of care, $\alpha G(\theta)/(1+t_f)$, internalizes the partner's valuation γ_m through $G(\theta)$, whereas in the semi-cooperative stage-2 problem each parent equates only their *private* benefit $\alpha\gamma_f/(1+t_f)$ to $\mu_f n/l_f$ (footnote to Section 4). Substituting the two hours conditions into the two care conditions gives the interior quality-care inputs in closed form,

$$1+t_f = \frac{\alpha\, G(\theta)\, c}{\phi(s) w_f\, n}, \qquad 1+t_m = \frac{(1-\alpha)\, G(\theta)\, c}{w_m\, n}, \tag{A6}$$

where $c = \tilde{Y}/\big(1+\theta\mu_f + (1-\theta)\mu_m + G(\theta)\big)$ and $\tilde{Y} = \phi(s) w_f\,(1+n-\tilde{t}_f n) + w_m\,(1+n-\tilde{t}_m n)$ is the household's full virtual income. The remaining eleven corner regimes replace $\tilde{Y}$ and the denominators by the corresponding virtual-endowment terms; a household-by-household check of the Karush–Kuhn–Tucker multipliers selects the unique valid regime, and the closed forms are validated against direct numerical maximization household by household (they attain at least the numerical optimum in 100% of 1,250 random draws spanning the estimation bounds).

Proof of Proposition 1, stated in the main text. Dividing the two expressions in (A6), the common factors $G(\theta)$, c, and n cancel, leaving $(1+t_f)/(1+t_m) = \frac{\alpha}{1-\alpha}\, w_m/(\phi(s) w_f)$, in which θ does not appear. Equivalently, θ enters the two interior care levels only through the common terms $G(\theta)$ and c, so it acts on t_f and t_m symmetrically. The wedge $\phi(s)$ instead enters only the wife's side, so it moves the relative price of her time and hence the ratio. $\square$

In the semi-cooperative benchmark the care ratio is instead $(1+t_f)/(1+t_m) = \frac{\alpha\gamma_f}{(1-\alpha)\gamma_m} \cdot l_f/l_m$, which depends on the equilibrium leisure (hours) allocation; holding the mean weight fixed, as the paper does, this too leaves the childcare reversals to the wedge. Section 4.4 of the main text discusses the corner-composition channel through which a wage-share–dependent weight can still bend the population care profiles.

J Solution of the semi-cooperative model

This appendix records the full closed-form solution of the two-stage game for completeness; the reader interested only in the mechanism can skip it. Backward induction yields twelve candidate regimes, indexed by which of the four non-negativity constraints bind: the second-stage childcare multipliers (η, λ) on (t_f, t_m) and the first-stage labor multipliers (τ, ν) on (h_f, h_m). Each $\{$*both, one, neither*$\}$-provides-care second-stage outcome combines with a $\{$*both, one, neither*$\}$-works first-stage outcome. The multiplier sign conditions select the *candidate* regimes but do not by themselves pin down the realized one: with log utility the reduced first-stage objective (the welfare index evaluated along the second-stage reaction functions) is concave on each care regime but only *piecewise* concave globally, because a spouse's reaction $t_i(h_i)$ has a kink where desired care reaches zero and the *partner's* valuation of t_i inherits it as a convex kink. Several regimes can therefore satisfy all their sign conditions at once, and the realized regime is the candidate attaining the highest first-stage welfare index $\theta u_f + (1-\theta)u_m$ (in the estimation sample up to four candidates pass the sign conditions simultaneously; the numerical implementation selects the welfare maximizer among them, validated household by household against direct numerical maximization of the reduced program). The single object that matters for the paper's argument is how the wedge $\phi(s)$ enters the interior solution (Regime 2a below), where it scales the wife's effective wage and so bends her market-hours and childcare responses back at parity. The remaining regimes are the corner cases.

The Lagrangian of the second stage problem is:

$$L_i = \delta_i \, \log\!\big(c\big) \; + \; \mu_i \, \log\!\big(1 - (t_i + \tilde{t}_i)\, n - h_i\big) \; + \; \gamma_i \, \log\!\big(n\,(1+t_i)^{\alpha}\,(1+t_{-i})^{1-\alpha}\big) \; + \; \eta\, t_i \; + \; \lambda\, t_{-i}$$

where η and λ are the Kuhn-Tucker multipliers associated to the non-negativity constraints on childcare time. The optimal solution satisfies $\eta t_f = \lambda t_m = 0$. The Nash Equilibrium of the second stage of the game can be in four regions, namely: $(\eta > 0, \lambda > 0), (\eta = 0, \lambda = 0), (\eta > 0, \lambda = 0), (\eta = 0, \lambda > 0)$.

The Lagrangian for the maximization of the first stage problem is

$$\begin{aligned} L = \theta\big(\delta_f \, \log(c) + \mu_f \, \log\!\big(1 - (t_f + \tilde{t}_f)\, n - h_f\big) + \gamma_f \, \log\!\big(n\,(1+t_f)^{\alpha}\,(1+t_m)^{1-\alpha}\big)\big) \\ +(1-\theta)\big(\delta_m \, \log(c) + \mu_m \, \log\!\big(1 - (t_m + \tilde{t}_m)\, n - h_m\big) \\ +\gamma_m \, \log\!\big(n\,(1+t_f)^{\alpha}\,(1+t_m)^{1-\alpha}\big)\big) + \tau\, h_f + \nu\, h_m, \end{aligned}$$

where τ and ν are the Kuhn-Tucker multipliers associated to the non-negativity constraints on working time. The optimal solution satisfies $\tau h_f = \nu h_m = 0$, with the exception of the case $h_f = h_m = 0$. Three work configurations for each of the four care regions give twelve time-allocation regimes, summarized below:

Regime	(η, λ)	(τ, ν)	Works	Provides quality care
1a	$> 0, > 0$	$= 0, = 0$	both	neither
1b	$> 0, > 0$	$= 0, > 0$	wife	neither
1c	$> 0, > 0$	$> 0, = 0$	husband	neither
2a	$= 0, = 0$	$= 0, = 0$	both	both
2b	$= 0, = 0$	$= 0, > 0$	wife	both
2c	$= 0, = 0$	$> 0, = 0$	husband	both
3a	$> 0, = 0$	$= 0, = 0$	both	husband
3b	$> 0, = 0$	$= 0, > 0$	wife	husband
3c	$> 0, = 0$	$> 0, = 0$	husband	husband
4a	$= 0, > 0$	$= 0, = 0$	both	wife
4b	$= 0, > 0$	$= 0, > 0$	wife	wife
4c	$= 0, > 0$	$> 0, = 0$	husband	wife

For compactness, the closed forms are reproduced here for two regimes only: regime 1b, a representative corner, and the interior regime 2a, the one the paper's argument turns on (it is where the wedge $\phi(s)$ scales the wife's effective wage and bends her market-hours and childcare responses at parity). The remaining ten closed forms are obtained by the same substitutions (each corner replaces the corresponding virtual-endowment terms in the stage-2 best responses and the reduced stage-1 objective) and are exported directly from the symbolic derivation in the replication code, where, as described above, every regime is validated household by household against direct numerical maximization of the reduced program.

Regime 1b: $\eta > 0, \lambda > 0, \tau = 0, \nu > 0$

$$h_m = 0$$

$$h_f = \frac{\big((\theta-1)n\tilde{t}_f - \theta + 1\big)\delta_m + \big(\theta - \theta n\tilde{t}_f\big)\,\delta_f}{\theta\,\mu_f + (1-\theta)\delta_m + \theta\delta_f}$$

$$t_m = 0$$

$$t_f = 0$$

Regime 2a: $\eta = 0, \lambda = 0, \tau = 0, \nu = 0$

$$h_m = \frac{\begin{aligned}&\Big((\theta-1)nw_f + (1-\theta)\phi n\Big)\tilde{t}_f + \Big((1-\theta)n - \theta + 1\Big)w_f + (\theta-1)\phi n + (\theta-1)\phi\Big)\,\mu_m \\ &+ \Big(\theta n w_m \tilde{t}_m + (-\theta n - \theta)w_m\Big)\,\mu_f \\ &+ \Big((1-\theta)nw_m\tilde{t}_m + \Big((\theta-1)n + \theta - 1\Big)w_m\Big)\,\delta_m \\ &+ \Big(\theta n w_m \tilde{t}_m + (-\theta n - \theta)w_m\Big)\,\delta_f \\ &+ \Big((\alpha - \alpha\theta)nw_m\tilde{t}_m + \Big((1-\alpha)\theta + \alpha - 1\Big)nw_f\tilde{t}_f + \Big((\alpha-1)\theta - \alpha + 1\Big)\phi n\tilde{t}_f \\ &\quad + (\alpha\theta - \alpha)nw_m + \Big(((\alpha-1)\theta - \alpha + 1)n + (\alpha-1)\theta - \alpha + 1\Big)w_f \\ &\quad + \Big((1-\alpha)\theta + \alpha - 1\Big)\phi n + \Big((1-\alpha)\theta + \alpha - 1\Big)\phi\Big)\gamma_m \\ &+ \Big(\alpha\theta n w_m\tilde{t}_m + \Big((\alpha-1)\theta n w_f + (1-\alpha)\theta\phi n\Big)\tilde{t}_f \\ &\quad + (-\alpha\theta n - \alpha\theta)w_m + \Big((1-\alpha)\theta n + (1-\alpha)\theta\Big)w_f + (\alpha-1)\theta\phi n + (\alpha-1)\theta\phi\Big)\gamma_f\end{aligned}}{\begin{aligned}(\theta-1)w_m\,\mu_m - \theta w_m\,\mu_f + (\theta-1)w_m\,\delta_m - \theta w_m\,\delta_f \\ +(\theta-1)w_m\,\gamma_m - \theta w_m\,\gamma_f\end{aligned}}$$

$$
h_f = -\frac{\begin{aligned}
&\Big((\theta-1)nw_f + (1-\theta)\phi n\Big)\tilde{t}_f + \big((1-\theta)n - \theta + 1\big)w_f + (\theta-1)\phi n + (\theta-1)\phi\Big)\,\mu_m \\
&+ \Big(\theta n w_m \tilde{t}_m + (-\theta n - \theta)w_m\Big)\,\mu_f \\
&+ \Big((\theta-1)nw_f + (1-\theta)\phi n\Big)\tilde{t}_f + \big((1-\theta)n - \theta + 1\big)w_f + (\theta-1)\phi n + (\theta-1)\phi\Big)\,\delta_m \\
&+ \Big((\theta\phi n - \theta n w_f)\tilde{t}_f + (\theta n + \theta)w_f - \theta\phi n - \theta\phi\Big)\,\delta_f \\
&+ \Big((\alpha - \alpha\theta)nw_m\tilde{t}_m + \big((1-\alpha)\theta + \alpha - 1\big)nw_f\tilde{t}_f + \big((\alpha-1)\theta - \alpha + 1\big)\phi n\tilde{t}_f \\
&\quad + (\alpha\theta - \alpha)nw_m + \big(((\alpha-1)\theta - \alpha + 1)n + (\alpha-1)\theta - \alpha + 1\big)w_f \\
&\quad + \big((1-\alpha)\theta + \alpha - 1\big)\phi n + \big((1-\alpha)\theta + \alpha - 1\big)\phi\Big)\gamma_m \\
&+ \Big(\alpha\theta n w_m\tilde{t}_m + \big((\alpha-1)\theta n w_f + (1-\alpha)\theta\phi n\big)\tilde{t}_f \\
&\quad + (-\alpha\theta n - \alpha\theta)w_m + \big((1-\alpha)\theta n + (1-\alpha)\theta\big)w_f + (\alpha-1)\theta\phi n + (\alpha-1)\theta\phi\Big)\gamma_f
\end{aligned}}{\begin{aligned}
&\Big((\theta-1)w_f + (1-\theta)\phi\Big)\mu_m + (\theta\phi - \theta w_f)\mu_f \\
&+ \Big((\theta-1)w_f + (1-\theta)\phi\Big)\delta_m + (\theta\phi - \theta w_f)\delta_f \\
&+ \Big((\theta-1)w_f + (1-\theta)\phi\Big)\gamma_m + (\theta\phi - \theta w_f)\gamma_f
\end{aligned}}
$$

$$
t_m = - \frac{
\begin{aligned}
&(\theta-1)nw_m\,\mu_m^2 \\
&\quad + \Bigl(-\theta n w_m\,\mu_f + (\theta-1)nw_m\,\delta_m - \theta n w_m\,\delta_f \\
&\qquad + \Bigl(((1-\alpha)\theta+\alpha-1)nw_m\tilde{t}_m + ((1-\alpha)\theta+\alpha-1)nw_f\tilde{t}_f \\
&\qquad\quad + \bigl((\alpha-1)\theta-\alpha+1\bigr)\phi n\tilde{t}_f + \bigl((\theta-1)n+(\alpha-1)\theta-\alpha+1\bigr)w_m \\
&\qquad\quad + \bigl(((\alpha-1)\theta-\alpha+1)n+(\alpha-1)\theta-\alpha+1\bigr)w_f \\
&\qquad\quad + \bigl((1-\alpha)\theta+\alpha-1\bigr)\phi n + \bigl((1-\alpha)\theta+\alpha-1\bigr)\phi\Bigr)\gamma_m - \theta n w_m\,\gamma_f\Bigr)\mu_m \\
&\quad + \alpha-1)\theta n w_m\,\gamma_m\,\mu_f \\
&\quad + \bigl((1-\alpha)\theta+\alpha-1\bigr)nw_m\,\gamma_m\,\delta_m + (\alpha-1)\theta n w_m\,\gamma_m\,\delta_f \\
&\quad + \Bigl(((\alpha^2-2\alpha+1)\theta-\alpha^2+2\alpha-1)nw_m\tilde{t}_m \\
&\qquad + ((\alpha^2-2\alpha+1)\theta-\alpha^2+2\alpha-1)nw_f\tilde{t}_f \\
&\qquad + \bigl((-\alpha^2+2\alpha-1)\theta+\alpha^2-2\alpha+1\bigr)\phi n\tilde{t}_f \\
&\qquad + \bigl(((\alpha-\alpha^2)\theta+\alpha^2-\alpha)n+(-\alpha^2+2\alpha-1)\theta+\alpha^2-2\alpha+1\bigr)w_m \\
&\qquad + \bigl((-\alpha^2+2\alpha-1)\theta n+(-\alpha^2+2\alpha-1)\theta\bigr)w_f \\
&\qquad + \bigl((\alpha^2-2\alpha+1)\theta-\alpha^2+2\alpha-1\bigr)\phi n + \bigl((\alpha^2-2\alpha+1)\theta-\alpha^2+2\alpha-1\bigr)\phi\Bigr)\gamma_m^2 \\
&\quad + \Bigl((-\alpha^2+2\alpha-1)\theta n w_m\tilde{t}_m + \bigl((-\alpha^2+2\alpha-1)\theta n w_f + (\alpha^2-2\alpha+1)\theta\phi n\bigr)\tilde{t}_f \\
&\qquad + \bigl((\alpha^2-\alpha)\theta n+(\alpha^2-2\alpha+1)\theta\bigr)w_m + \bigl((\alpha^2-2\alpha+1)\theta n+(\alpha^2-2\alpha+1)\theta\bigr)w_f \\
&\qquad + (-\alpha^2+2\alpha-1)\theta\phi n + (-\alpha^2+2\alpha-1)\theta\phi\Bigr)\gamma_f\gamma_m
\end{aligned}
}{
\begin{aligned}
&(\theta-1)nw_m\,\mu_m^2 \\
&\quad + \Bigl(-\theta n w_m\,\mu_f + (\theta-1)nw_m\,\delta_m - \theta n w_m\,\delta_f \\
&\qquad + \bigl((2-2\alpha)\theta+2\alpha-2\bigr)nw_m\,\gamma_m - \theta n w_m\,\gamma_f\Bigr)\mu_m \\
&\quad + \alpha-1)\theta n w_m\,\gamma_m\,\mu_f \\
&\quad + \bigl((1-\alpha)\theta+\alpha-1\bigr)nw_m\,\gamma_m\,\delta_m + (\alpha-1)\theta n w_m\,\gamma_m\,\delta_f \\
&\quad + \bigl((\alpha^2-2\alpha+1)\theta-\alpha^2+2\alpha-1\bigr)nw_m\,\gamma_m^2 \\
&\quad + \bigl(-\alpha^2+2\alpha-1\bigr)\theta n w_m\,\gamma_f\gamma_m
\end{aligned}
}
$$

$$t_f = -\frac{\begin{aligned}&\big((\theta-1)nw_f+(1-\theta)\phi n\big)\mu_f\,\mu_m+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_f\mu_m\\&+\big(\theta\phi n-\theta nw_f\big)\mu_f^2\\&+\big((\theta-1)nw_f+(1-\theta)\phi n\big)\delta_m+\big(\theta\phi n-\theta nw_f\big)\delta_f\\&+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_m\\&+\Big(-\alpha\theta nw_m\tilde{t}_m+(\alpha\theta\phi n-\alpha\theta nw_f)\tilde{t}_f+\alpha\theta w_m+(\alpha\theta-\alpha\theta n)w_f+\alpha\theta\phi n-\alpha\theta\phi\Big)\gamma_f\\&+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_f\delta_m+\big(\alpha\theta\phi n-\alpha\theta nw_f\big)\gamma_f\delta_f\\&+\Big((\alpha^2\theta-\alpha^2)nw_m\tilde{t}_m+\big((\alpha^2\theta-\alpha^2)nw_f\\&+(\alpha^2-\alpha^2\theta)\phi n\big)\tilde{t}_f+(\alpha^2-\alpha^2\theta)(w_m+w_f)+(\alpha^2\theta-\alpha^2)\phi\Big)\gamma_f\gamma_m\\&+\Big(-\alpha^2\theta nw_m\tilde{t}_m+(\alpha^2\theta\phi n-\alpha^2\theta nw_f)\tilde{t}_f+\alpha^2\theta(w_m+w_f)-\alpha^2\theta\phi\Big)\gamma_f^2\end{aligned}}{\begin{aligned}&\big((\theta-1)nw_f+(1-\theta)\phi n\big)\mu_f\,\mu_m+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_f\mu_m\\&+\big(\theta\phi n-\theta nw_f\big)\mu_f^2\\&+\big((\theta-1)nw_f+(1-\theta)\phi n\big)\delta_m+\big(\theta\phi n-\theta nw_f\big)\delta_f\\&+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_m\\&+\big(2\alpha\theta\phi n-2\alpha\theta nw_f\big)\gamma_f\\&+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_f\delta_m+\big(\alpha\theta\phi n-\alpha\theta nw_f\big)\gamma_f\delta_f\\&+\big((\alpha\theta-\alpha)nw_f+(\alpha-\alpha\theta)\phi n\big)\gamma_f\gamma_m+\big(\alpha\theta\phi n-\alpha\theta nw_f\big)\gamma_f^2\end{aligned}}$$

This completes the closed forms the estimation uses directly; the other ten regimes live in the replication code.

References

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Adachi, D., Kawaguchi, D., and Saito, Y. U. (2024). Robots and employment: Evidence from Japan, 1978–2017. *Journal of Labor Economics*, 42(2):591–634.

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Anelli, M., Giuntella, O., and Stella, L. (2024). Robots, marriageable men, family, and fertility. *Journal of Human Resources*, 59(2):443–469.

Autor, D., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–78.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–68.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Besedeš, T., Lee, S. H., and Yang, T. (2021). Trade liberalization and gender gaps in local labor market outcomes: Dimensions of adjustment in the United States. *Journal of Economic Behavior & Organization*, 183:574–588.

Dauth, W., Findeisen, S., Suedekum, J., and Woessner, N. (2021). The adjustment of labor markets to robots. *Journal of the European Economic Association*, 19(6):3104–3153.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Halla, M., Schmieder, J., and Weber, A. (2020). Job displacement, family dynamics, and spousal labor supply. *American Economic Journal: Applied Economics*, 12(4):253–87.

Heggeness, M. and Murray-Close, M. (2019). Manning up and womaning down: How husbands and wives report earnings when she earns more. Opportunity and Inclusive Growth Institute Working Papers 28, Federal Reserve Bank of Minneapolis.

ISSP Research Group (2016). International Social Survey Programme: Family and Changing Gender Roles IV – ISSP 2012. GESIS Datenarchiv, Köln. Available online at https://doi.org/10.4232/1.12661.

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Technical Report DAMTP 2009/NA06, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.

Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534.

Rosenberg, S. (2021). Revisiting the Breadwinner Norm: The effect of potential relative earnings on married women's labor supply. Manuscript, ECARES, Université Libre de Bruxelles.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rodgers, R., and Schouweiler, M. (2024). IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS.