Additional Results Not for Publication

Cognitive Skills and the Timing of First Births

Agustín Díaz Casanueva[*]

September 26, 2026

This document collects results that support the paper and its Online Appendix but are not part of either: additional descriptive tables, further robustness of the reduced-form gradient, auxiliary-regression coefficient tables, secondary counterfactual variants and the computational documentation. Sections are numbered S.1, S.2, …; tables and figures S.1, S.2, … All numbers are produced by the same pipeline as the paper. References to "Section X", "Table X" or "Online Appendix OA.X" without an S prefix point to the paper and its Online Appendix.

Contents

[*]Central Bank of Chile, adiaz@bcentral.cl.

Table S.1. Educational Attainment by Cognitive Ability Quartile

	Cognitive Ability (AFQT) Quartile				
Education outcome	Q1 (Low)	Q2	Q3	Q4 (High)	All
HS dropout	29	9	2	1	10
HS graduate	68	82	77	48	69
College attendance	11	25	42	69	37
College graduate	3	9	20	51	21
Women	1,455	1,455	1,455	1,455	5,820

Notes: Estimation panel: the same women as Table OA.3 and the birth moments of Section 4. Education is the highest degree completed; college attendance is attendance between ages 18 and 22. Entries are column percentages, the share of women in each AFQT quartile with the indicated outcome. The college-attendance row is the attendance moment the model targets (11.5, 24.7, 42.1 and 68.7 percent), computed on the same women.

Appendices

S.1 Additional Descriptive Statistics

This section reports descriptive tables that supplement Section 2.4 of the paper and Online Appendix OA.2: the distribution of education by ability quartile and the earnings of husbands by the wife's education and the timing of her first birth relative to marriage. Variables and samples are defined in Online Appendix OA.1.

Table S.2 reports average husbands' annual wages by the woman's education, age at first birth, and whether the first birth occurs out of wedlock, conditional on marrying. In most cells, women with an out-of-wedlock first birth have lower-earning husbands; the exception is teen mothers without a high-school degree. The spousal-earnings differential is largest for high-school graduates with a first birth at 22–29—about \$17,700 (60,614 vs. 42,945). These spousal-earnings penalties represent an additional cost of imperfect fertility control that the model incorporates through the marriage market.

Table S.2. Average Husband Wage by Education and Women's Childbearing Status at Marriage

	HS Dropout		HS Graduate		College Graduate	
Age at first birth	Out-wed.	In-wed.	Out-wed.	In-wed.	Out-wed.	In-wed.
14–17	41,135	36,728	47,013	53,855	—	—
18–21	37,672	44,185	44,588	51,900	—	64,394
22–29	—	36,737	42,945	60,614	—	82,254

Notes: Mean annual wage-and-salary income of the husband (2016 dollars) over married woman-years, by the woman's final education, her age at first birth, and whether her first marriage came after (out of wedlock) or before (in wedlock) the first birth. Spouse-years as in the husband earnings regression (Table S.17): at least 26 weeks worked, more than 20 hours a week and income of at least $10,500; income refers to the calendar year before the interview and is deflated with that year's CPI. —: fewer than 30 women in the cell.

S.2 The Wantedness Status of First Births

Table S.3 reports, by AFQT quartile, the distribution of first births across the five contraceptive-status groups at conception that Online Appendix OA.3 describes: the method in use failed; the woman had stopped using a method, not in order to conceive; she had used none since the previous interview and was not trying to conceive; she had stopped in order to conceive; and she was using nothing because she wanted a birth. Panel C is the source of the composition cells of panel C of Table OA.36 in the paper.

Table S.3. Contraceptive status and wantedness of the first birth, by AFQT quartile

	Q1	Q2	Q3	Q4	All
A. Status at the conception of the first birth (% of first births)					
Using a method, which failed	8.5	10.6	15.6	16.6	12.6
Had stopped, not to conceive	9.9	10.2	9.2	7.6	9.3
Never used, not wanting a birth	39.0	33.4	21.9	12.2	27.4
Had stopped in order to conceive	13.3	18.7	23.1	28.1	20.4
Not using, wanted a birth	29.3	27.2	30.1	35.4	30.3
B. Mistimed or unwanted (% of first births), by status at conception					
Mistimed or unwanted	46.5	43.9	38.8	30.5	40.4
method failed	6.9	8.9	14.2	14.3	10.8
had stopped, not to conceive	8.6	8.7	7.4	6.8	7.9
never used, not wanting	30.9	26.3	17.2	9.3	21.6
No method (% of mistimed or unwanted)	85.1	79.7	63.4	53.0	73.2
C. Mistimed or unwanted (% of first births), by age at first birth					
14–17	62.7	68.9	69.7	69.6	66.0
18–21	45.6	47.8	51.6	56.4	49.2
22–25	29.0	31.4	32.9	33.0	31.8
26–41	25.6	22.6	17.8	14.3	18.1
Women	1,143	1,079	959	943	4,124

Notes: Sample: women in the structural sample with a first birth at ages 14–41 whose first-birth pregnancy record is matched and has a classifiable contraceptive status. Status at conception is from Q9-78 to Q9-80 on that record; "mistimed or unwanted" is Q9-81 answered "not at that time" or "not at all". Panel A: percent of first births. Panel B: percent of first births that were mistimed or unwanted, in total and by status at conception; its first row is the targeted moment, which in addition includes the ten women whose record carries a wantedness answer but no classifiable status. The last row of Panel B is the percent of mistimed or unwanted first births conceived with no method in use. Panel C: the same share by age at first birth; the teenage cell of the top quartile holds 46 births.

S.3 Further Robustness of the Reduced-Form Gradient

This section extends Online Appendix OA.5. Tables S.4–S.7 are the within-education, unconditional and unintended-birth versions of the specification ladder of Online Appendix OA.5.1,

Table S.4. Within-Education Birth Probability by AFQT Quartile: Baseline vs. Full Controls

	No HS diploma		HS / some college		College	
	(1) Baseline	(2) Full	(3) Baseline	(4) Full	(5) Baseline	(6) Full
AFQT Q2	-0.019*	0.001	-0.019***	-0.013**	0.056**	0.056**
	(0.010)	(0.010)	(0.005)	(0.005)	(0.028)	(0.026)
AFQT Q3	-0.034**	-0.012	-0.021***	-0.006	0.017	0.024
	(0.015)	(0.016)	(0.005)	(0.006)	(0.025)	(0.022)
AFQT Q4	-0.045**	-0.004	-0.029***	-0.004	0.023	0.035
	(0.018)	(0.021)	(0.006)	(0.007)	(0.024)	(0.022)
Observations	5,251	5,251	22,342	22,342	4,498	4,498

Notes: Sample and outcome as in Table OA.11, Panel A. "Baseline" is Table OA.11. "Full" adds enrollment status, age-adjusted AFQT quartiles, race and family background; unweighted. SE clustered by respondent. Reference: AFQT Q1. The unintended-birth version is in Table S.5.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

whose pooled and by-method versions are in the Online Appendix; the text there discusses all six. Figure S.1 plots the within-method birth probabilities by quartile and method group. The remaining subsections report the relationship between cognitive ability and family background and the timing of the AFQT relative to first births and attrition.

S.3.1 Additional Controls and Sampling Weights: Further Specifications

S.3.2 Cognitive Ability and Family Background

Table S.8 shows how AFQT relates to parents' schooling, 1979 family income, 1979 poverty and race. The gradient in family background is steep: 28% of mothers of lowest-quartile women completed high school against 82% in the top quartile, and 48% of lowest-quartile families were below the poverty line in 1979 against 12%. Family background and race together explain a fraction $R^2 = 0.37$ of the variance in AFQT (0.28 excluding race). This is the premise of the premarket-skill literature: measured cognitive ability at ages 15–23 is itself an outcome of family inputs (Carneiro and Heckman, 2003; Heckman et al., 2006; Neal and Johnson, 1996), so the question is not whether AFQT is correlated with background—it

Table S.5. Within-Education Unintended-Birth Probability by AFQT Quartile: Baseline vs. Full Controls

	No HS diploma		HS / some college		College	
	(1) Baseline	(2) Full	(3) Baseline	(4) Full	(5) Baseline	(6) Full
AFQT Q2	0.005	0.014**	-0.010***	-0.005	0.003	0.002
	(0.007)	(0.007)	(0.003)	(0.003)	(0.024)	(0.022)
AFQT Q3	-0.005	0.010	-0.009***	0.001	-0.021	-0.009
	(0.011)	(0.012)	(0.003)	(0.004)	(0.021)	(0.019)
AFQT Q4	-0.020*	0.010	-0.012***	0.002	-0.016	-0.004
	(0.011)	(0.017)	(0.004)	(0.004)	(0.021)	(0.019)
Observations	5,149	5,149	22,023	22,023	4,476	4,476

Notes: As Table S.4, outcome = birth reported as unwanted or mistimed (Table OA.11, Panel B).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S.6. Birth Probability Unconditional on Method: Additional Controls and Sampling Weights (LPM)

	(1) Previous	(2) Baseline	(3) + Enrol.	(4) Age-adj.	(5) + Family	(6) + Race	(7) + Both
Panel A: Unweighted							
AFQT Q2	-0.014***	-0.019***	-0.018***	-0.017***	-0.017***	-0.009**	-0.010**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
AFQT Q3	-0.017***	-0.024***	-0.022***	-0.022***	-0.019***	-0.006	-0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
AFQT Q4	-0.022***	-0.030***	-0.025***	-0.024***	-0.020***	-0.002	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Panel B: NLSY79 sampling weights							
AFQT Q2	-0.006	-0.014**	-0.014**	-0.013**	-0.012**	-0.005	-0.006
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
AFQT Q3	-0.004	-0.015***	-0.013**	-0.013**	-0.011*	-0.002	-0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
AFQT Q4	-0.008	-0.019***	-0.014**	-0.013**	-0.010	0.002	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Test weighted = unweighted, p-value	0.001	0.003	0.002	0.004	0.034	0.462	0.619
Observations	32,091	32,091	32,091	32,091	32,091	32,091	32,091

Notes: Sample and outcome as in Table OA.24, column (1). Columns and panels as in Table OA.13, without method-group FE. SE clustered by respondent. Reference: AFQT Q1.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S.7. Unintended-Birth Probability Unconditional on Method: Additional Controls and Sampling Weights (LPM)

	(1) Previous	(2) Baseline	(3) + Enrol.	(4) Age-adj.	(5) + Family	(6) + Race	(7) + Both
Panel A: Unweighted							
AFQT Q2	-0.006**	-0.006**	-0.006*	-0.006*	-0.005*	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
AFQT Q3	-0.010***	-0.009***	-0.008***	-0.007**	-0.005*	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
AFQT Q4	-0.011***	-0.011***	-0.009***	-0.008**	-0.006*	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Panel B: NLSY79 sampling weights							
AFQT Q2	-0.003	-0.003	-0.003	-0.002	-0.001	0.003	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
AFQT Q3	-0.006*	-0.006*	-0.006	-0.003	-0.001	0.004	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
AFQT Q4	-0.005	-0.006*	-0.005	-0.002	0.001	0.007*	0.008**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Test weighted = unweighted, p-value	0.080	0.258	0.231	0.054	0.034	0.262	0.214
Observations	31,648	31,648	31,648	31,648	31,648	31,648	31,648

Notes: Outcome: birth reported as unwanted or mistimed (Table OA.24, column (2)). Otherwise as Table S.6.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

is—but whether background can substitute for it in predicting fertility.

Tables S.9 and S.10 answer that question in three steps. Column (2) uses only the background variables and no AFQT; column (3) uses both; column (4) adds welfare receipt and poverty status in the wave year as contemporaneous proxies, with the caveat that they are themselves outcomes of early births (Medicaid and health-insurance items are asked only from 1989 and are not used). Parents' schooling does not predict births conditional on the baseline controls—mother's and father's high-school completion carry coefficients of 0.002 and -0.004—and neither does it move the AFQT coefficient. Log family income in 1979 does predict births, at 0.010 (s.e. 0.002), but its sign runs *against* the ability gradient: conditional on ability, women from higher-income families have higher birth rates in a given wave, so controlling for income makes the AFQT coefficient larger, not smaller. The Gelbach (2016) decomposition puts a number on this—income's share of the movement in the top-quartile coefficient is -20%—and the background block as a whole is not what absorbs the

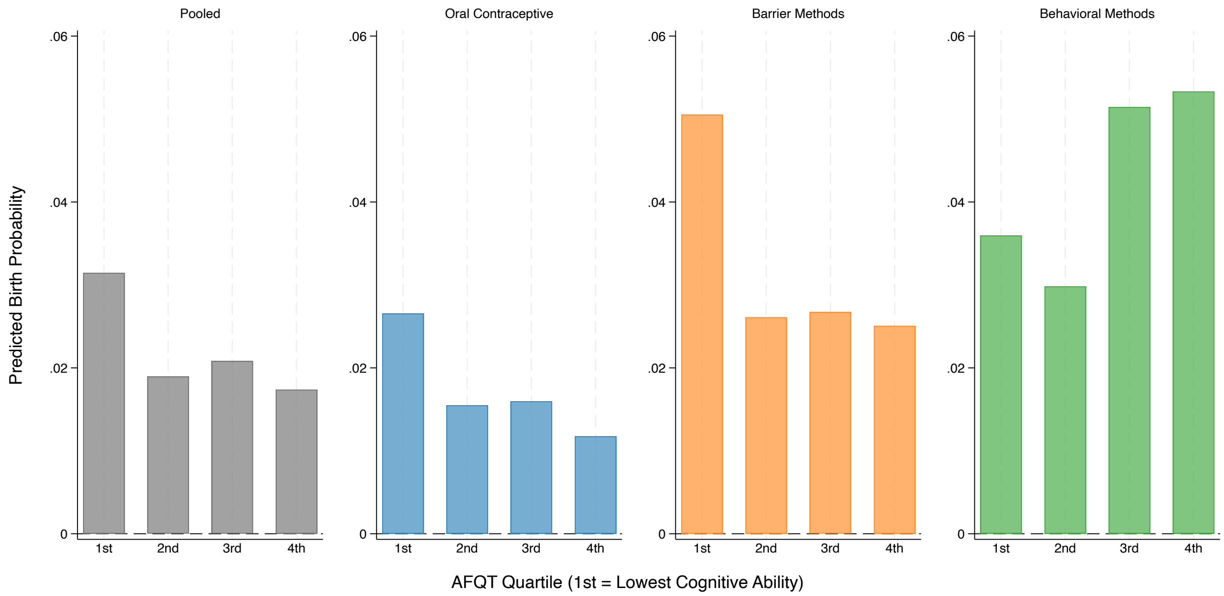


Figure S.1. LPM-predicted birth probability by AFQT quartile within each method group. Bars plot LPM predicted probabilities evaluated at each AFQT quartile, averaged over the sample distribution of controls (education group, parity, age and marital status at the previous interview). Pooled specification adds method-group fixed effects. Denominator: NLSY79 women reporting use of method n at the snapshot-year interview; sample pools seven cohorts ($t \in \{1982, 1984, 1985, 1986, 1988, 1990, 1992\}$), ages 18–35. Outcome: a live birth in a fixed 12-month window (July t to June $t+1$) that the woman reports she had not wanted then, under the classification of Online Appendix OA.1.7. Snapshot-year FEs included in all specifications.

gradient. What moves the AFQT coefficient is race: in the unconditional sample the top-quartile coefficient goes from -0.030 to -0.012 (61% absorbed) once both blocks are added, and within method from -0.014 to -0.005 (66%), while family background on its own leaves both where they were. Online Appendix OA.5.2 reports the two blocks separately and gives the Gelbach (2016) decomposition of the unconditional sample (Table OA.16); Table S.11 here is its within-method counterpart. Both assign more than all of the change to the Black and Hispanic indicators, essentially none to parental schooling, and a negative share to 1979 income. This is not an SES explanation of the gradient; it is the mechanical consequence of the fact that 47% of lowest-quartile women are Black and 7% of top-quartile women are. A race indicator removes the between-group component of AFQT by construction. Whether the ability gradient is a race gradient in disguise is a different question, and it has to be answered the way the paper answers the same question for education—within groups.

Table S.8. Family Background by AFQT Quartile

	Q1 (lowest)	Q2	Q3	Q4 (highest)	All
Mother's highest grade	8.95	10.25	11.37	12.53	10.82
Father's highest grade	8.47	9.88	11.45	13.13	10.89
Mother completed HS	0.28	0.44	0.65	0.82	0.55
Father completed HS	0.31	0.44	0.64	0.81	0.57
Family income 1979 (thousands of dollars)	9.78	13.37	15.36	19.85	14.61
Family in poverty 1979	0.48	0.30	0.17	0.12	0.27
Black	0.47	0.31	0.16	0.07	0.25
Hispanic	0.25	0.18	0.12	0.07	0.16
Mean AFQT percentile	8.4	27.7	51.0	81.2	42.2
Women	1,473	1,483	1,493	1,490	5,939

Notes: NLSY79 women with a valid AFQT score; unweighted means. Parents' highest grade and family income are from the 1979 interview, when respondents were aged 14–22 and most lived with their parents; family income is total net family income (nominal dollars, truncated by the NLSY79) and includes parental income for dependent respondents.

Appendix OA.5.6 does so.

S.3.3 The Failure Rate the Method Mix Alone Would Predict

Table S.12 assigns each woman the typical-use failure rate of her most effective reported method (Trussell, 2011) and averages by AFQT quartile; the resulting predicted Q1/Q4 failure ratio is compared with the observed birth-rate ratio in Online Appendix OA.6.1.

S.3.4 Men and Marital Sorting

Table S.14 reports the ability gradient in the timing of first parenthood for men alongside women; women's quartiles are the paper's (Appendix OA.1.3), men's are drawn within sex. The gradient is present for men and about two-thirds as steep: 37.5% of lowest-quartile men have a first child by 22 against 10.7% of the highest quartile (a gap of 26.8 points, versus 43.0 points for women), and the mean age at first child rises from 23.4 to 27.9 across quartiles (20.9 to 26.2 for women). In the regression version (Table S.15) the quartile $\times$ male interactions are positive, so the male gradient is significantly flatter at ages 22 and 25 and indistinguishable from the female one by 30. Men's reports of children fathered are less complete than women's (Rendall et al., 1999), and under-reporting is concentrated among young, unmarried fathers, which compresses the male gradient at early ages; the comparison

Table S.9. The Ability Gradient With and Without Family-Background Controls: Unconditional Sample (LPM)

	(1) AFQT only	(2) SES only	(3) AFQT + SES	(4) + contemp. SES
AFQT Q2	-0.019***		-0.012**	-0.006
	(0.005)		(0.005)	(0.005)
AFQT Q3	-0.024***		-0.011**	-0.004
	(0.005)		(0.005)	(0.005)
AFQT Q4	-0.030***		-0.012**	-0.005
	(0.005)		(0.005)	(0.005)
Mother completed HS		0.002	0.002	0.002
		(0.004)	(0.004)	(0.004)
Father completed HS		-0.005	-0.004	-0.003
		(0.004)	(0.004)	(0.004)
Log family income 1979		0.010***	0.010***	0.010***
		(0.002)	(0.002)	(0.002)
In poverty 1979		0.009*	0.008	0.005
		(0.005)	(0.005)	(0.005)
Black		0.036***	0.033***	0.030***
		(0.004)	(0.004)	(0.004)
Hispanic		0.034***	0.032***	0.033***
		(0.004)	(0.005)	(0.005)
Welfare receipt at wave				0.055***
				(0.006)
In poverty at wave				-0.007
				(0.004)
Observations	32,091	32,091	32,091	32,091

Notes: Sample, outcome (any birth) and baseline controls as in Table OA.24, column (1). Column (1): AFQT quartiles only; (2): background only (mother and father completed high school, log 1979 family income, 1979 poverty, race; missing indicators included, not shown); (3): both; (4): adds welfare receipt and poverty status in the wave year. SE clustered by respondent. Reference: AFQT Q1; non-Black non-Hispanic.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

at 30, where the two gradients coincide, is the more reliable one.

Marital sorting on ability is the other half of the same question: if low-ability women marry high-ability men, household resources and preferences after marriage differ from what the woman's own characteristics imply. Table S.13 reports partner earnings of married women aged 28–40 by the woman's education and AFQT quartile. Sorting is strong and it works *with* the woman's ability rather than against it: within the high-school group, partners

Table S.10. The Ability Gradient With and Without Family-Background Controls: Within Method (LPM)

	(1) AFQT only	(2) SES only	(3) AFQT + SES	(4) + contemp. SES
AFQT Q2	-0.013***		-0.008*	-0.007
	(0.004)		(0.004)	(0.004)
AFQT Q3	-0.011**		-0.003	-0.002
	(0.004)		(0.005)	(0.005)
AFQT Q4	-0.014***		-0.005	-0.003
	(0.004)		(0.005)	(0.005)
Mother completed HS		0.004	0.004	0.004
		(0.003)	(0.003)	(0.003)
Father completed HS		-0.003	-0.003	-0.002
		(0.003)	(0.003)	(0.003)
Log family income 1979		0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)
In poverty 1979		0.004	0.003	0.003
		(0.005)	(0.005)	(0.005)
Black		0.022***	0.021***	0.021***
		(0.003)	(0.004)	(0.004)
Hispanic		0.006*	0.005	0.006
		(0.004)	(0.004)	(0.004)
Welfare receipt at wave				0.021***
				(0.006)
In poverty at wave				-0.006
				(0.004)
Observations	16,087	16,087	16,087	16,087

Notes: Sample, outcome and baseline controls as in Table OA.12, column (1). Columns as in Table S.9. SE clustered by respondent. Reference: AFQT Q1; non-Black non-Hispanic.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of top-quartile women earn 38 log points more than partners of bottom-quartile women; within college, 47 log points; among dropouts the point estimates are larger still (50–65 log points) but less precise. Marriage itself is also sorted—43% of lowest-quartile women are married at ages 28–40 against 70% of the highest. In the model, the marriage market enters through the meeting probability $\mu(e_t, t)$, the partner earnings process of Appendix OA.10, which conditions on the woman's education and on whether she married before her first birth, and the marital surplus. The partner earnings process does not condition on the

Table S.11. Gelbach Decomposition of the Change in the Q4 Coefficient (Within Method)

	Contribution to Q4 gap	Share of gap
Mother completed HS	0.0017	-18.7%
Father completed HS	-0.0011	11.7%
Log family income 1979	0.0014	-15.5%
In poverty 1979	-0.0008	9.1%
Black	-0.0100	107.8%
Hispanic	-0.0008	8.4%
Family background (four rows above)	0.0012	-13.3%
Race and ethnicity	-0.0107	116.2%
Total (baseline minus controlled)	-0.0095	100%

Notes: As in Table OA.16 of the Online Appendix, estimated on the within-method sample: each row is $\hat{\gamma}_k \hat{\beta}_k$, where $\hat{\gamma}_k$ is the Q4 coefficient in a regression of background variable k on the AFQT quartile dummies and baseline controls, and $\hat{\beta}_k$ its coefficient in column (3) of Table S.10 (Gelbach, 2016). Rows sum to the difference between the Q4 coefficients in columns (1) and (3) up to the missing-indicator terms.

Table S.12. Failure Rate Predicted by the Method Mix vs. Observed Birth Rate, by AFQT Quartile

	Predicted from method mix		Observed (NLSY79)	
AFQT quartile	Typical use (%)	Perfect use (%)	12-month birth rate (%)	Woman-waves
Q1 (lowest)	12.60	2.93	3.59	3,289
Q2	12.09	2.02	2.13	3,936
Q3	11.92	1.74	1.95	4,351
Q4 (highest)	11.92	1.89	1.32	4,548
Ratio Q1 / Q4	1.06	1.55	2.72	

Notes: Predicted columns: each woman-wave using a pill, barrier or behavioral method is assigned the first-year failure rate of her most effective reported method (Trussell, 2011); the table reports the quartile mean. Observed: 12-month birth rate among the same population (Table OA.12 sample). Levels are not comparable across the two panels (annual failure among all users vs. live births in a fixed window among women not trying to conceive); the Q1/Q4 ratios are.

woman's ability within education, so the model understates the marriage-market return to ability for high-ability women; this makes the opportunity-cost channel, if anything, weaker in the model than in the data and works against attributing the fertility gradient to fertility control. What sorting cannot do is explain the timing gradient at ages 14–21, which precedes marriage for most women in the sample: 63% of lowest-quartile women have a first birth by age 21 (Table OA.22) and only 48% of lowest-quartile women are married at their first birth (Table OA.3).

Table S.13. Partner Earnings by the Woman's Education and AFQT Quartile

	Mean partner earnings (2016 $)				Log partner earnings, coef. on quartile			
Woman's education	Q1	Q2	Q3	Q4	Q2	Q3	Q4	Woman-years
Dropout	30,256	46,363	46,900	49,130	0.50*** (0.10)	0.65*** (0.11)	0.63*** (0.21)	1,129
High school	41,719	49,898	57,652	59,467	0.22*** (0.04)	0.34*** (0.04)	0.38*** (0.04)	11,508
College	49,833	57,367	65,599	80,994	0.15 (0.12)	0.28*** (0.11)	0.47*** (0.10)	5,124

Notes: Married NLSY79 women aged 28–40 with a valid AFQT score and positive partner wage-and-salary income, waves 1985–2000; woman-year observations; quartiles at the estimation-sample cut-points (Appendix OA.1.3). Means in 2016 dollars, deflated by the CPI series the earnings process uses. Regression columns: log partner earnings on AFQT quartile dummies, year fixed effects and the woman's age, estimated within education group; SE clustered by respondent in parentheses. Reference: AFQT Q1. Education is the woman's highest grade completed over the panel.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S.14. First Parenthood by AFQT Quartile: Women and Men

	Women					Men				
	Q1	Q2	Q3	Q4	Q1 minus Q4	Q1	Q2	Q3	Q4	Q1 minus Q4
First child by age 20	0.493	0.348	0.216	0.120	0.373	0.222	0.163	0.104	0.052	0.170
First child by age 22	0.633	0.488	0.336	0.203	0.430	0.375	0.288	0.203	0.107	0.268
First child by age 25	0.752	0.642	0.474	0.336	0.417	0.532	0.463	0.356	0.217	0.315
First child by age 30	0.827	0.753	0.640	0.555	0.272	0.673	0.614	0.533	0.431	0.242
First child by age 40	0.876	0.824	0.733	0.717	0.159	0.738	0.694	0.663	0.621	0.117
Mean age at first child (parents by 40)	20.9	22.5	24.1	26.2	-5.2	23.4	24.2	25.8	27.9	-4.4
Respondents	1,473	1,483	1,493	1,490		1,494	1,494	1,494	1,493	

Notes: NLSY79 respondents with a valid AFQT score, the 5,939 women and the men; women's quartiles use the estimation-sample cut-points (Appendix OA.1.3) and the 119 women the model excludes are kept, since the comparison with men does not use the model; men's quartiles are drawn within sex. First child from the NLSY79 fertility histories (year of birth of the first child); men's histories are self-reported children fathered. The regression version with quartile × sex interactions is in Table S.15.

Table S.15. First Child by a Given Age on AFQT Quartile, Sex, and Their Interaction (LPM)

	(1) First child by 22	(2) First child by 25	(3) First child by 30
AFQT Q2	-0.136***	-0.102***	-0.067***
	(0.018)	(0.017)	(0.015)
AFQT Q3	-0.280***	-0.262***	-0.173***
	(0.018)	(0.018)	(0.017)
AFQT Q4	-0.407***	-0.396***	-0.255***
	(0.017)	(0.018)	(0.017)
Male	-0.257***	-0.218***	-0.152***
	(0.018)	(0.017)	(0.016)
Q2 x Male	0.056**	0.037	0.013
	(0.025)	(0.025)	(0.023)
Q3 x Male	0.125***	0.100***	0.045*
	(0.024)	(0.025)	(0.024)
Q4 x Male	0.162***	0.099***	0.027
	(0.022)	(0.024)	(0.024)
Respondents	11,914	11,914	11,914

Notes: Controls: race and birth-year dummies. Robust SE. Reference: AFQT Q1, women.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

S.4 Wage-Process Estimates and Retirement Income

This section reports the estimated coefficients of the earnings regressions specified in Online Appendix OA.10 and the construction of the retirement-income process.

S.4.1 Estimated parameters

Interpreting the ability gradient. Table S.16 shows the estimated parameters of the women's income process. The estimates reveal substantial wage premia associated with cognitive ability. The baseline ability effect (e.g., Q4 vs. Q1 for high-school dropouts: $2,954) reflects both direct productivity differences and indirect effects through sorting into higher-paying jobs. The college-ability interactions are particularly large (the college-by-ability interactions for Q2–Q4 are $5,700–$7,400, on top of the ability and college main effects), consistent with complementarity between cognitive ability and advanced education in the labor market.

Interpreting the education coefficients. Table S.17 shows the estimated parameters of the husband's income process. The negative main effects for wife's education (-$28,456 for HS grad, -$116,190 for college) reflect the specification's interaction structure. These coefficients capture husband earnings when the wife marries after first birth (the baseline) and when the wife is at age zero. The large positive age-education interactions ($1,227/year for HS grad wives, $6,109/year for college wives) indicate that the husbands of more-educated women have steeper earnings growth over the wife's life cycle. At typical ages of the wife (e.g., 35), college-educated women's husbands earn substantially more than dropout women's husbands, consistent with positive assortative mating. The marriage-timing interactions show that women who marry before first birth (a marker of deliberate family formation) have higher-earning husbands, with this premium larger for more-educated women.

S.4.2 Retirement income process

The NLSY wage-and-salary measures do not capture the older-age components of retirement resources (Social Security benefits, employer pensions, and other transfers) because

Table S.16. Women's Earnings Process Estimates

Variable	Coef. (SE)	Variable	Coef. (SE)
Experience		*Education × ability*	
Cumulative experience	968***	HS grad × Q2	77
	(33)		(630)
		HS grad × Q3	1558
Education (base: HSD)			(963)
HS graduate	1061***	HS grad × Q4	514
	(395)		(1788)
College graduate	3725***	College grad × Q2	7379***
	(1123)		(1392)
		College grad × Q3	6523***
Education × experience			(1510)
HS grad × exp	232***	College grad × Q4	5712***
	(35)		(2104)
College grad × exp	758***		
	(96)	*Educ × ability × exp*	
		HS grad × Q2 × exp	141**
Ability (base: Q1)			(60)
Ability Q2	1228**	HS grad × Q3 × exp	66
	(562)		(80)
Ability Q3	1419	HS grad × Q4 × exp	46
	(916)		(123)
Ability Q4	2954*	College grad × Q2 × exp	-217*
	(1756)		(116)
		College grad × Q3 × exp	181
Ability × experience			(127)
Q2 × exp	-31	College grad × Q4 × exp	467***
	(55)		(156)
Q3 × exp	86		
	(76)	*Age profile*	
Q4 × exp	321***	Age	1406***
	(120)		(119)
		Age^2	-22***
			(2)
		Constant	-5994***
			(2009)
Observations			91,410
Adjusted R^2			0.294

Notes: Robust standard errors in parentheses; year fixed effects included. Dependent variable is annual wage-and-salary income in 2016 dollars. Baselines: high-school dropout (HSD) and ability quartile 1 (Q1). Coefficients and standard errors rounded to nearest dollar. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table S.17. Husband/Partner Earnings Process Estimates

Variable	Coef. (SE)	Variable	Coef. (SE)
Wife's education (base: HSD)		*Age profile*	
HS graduate	-28456***	Age	731
	(7054)		(498)
College graduate	-116190***	HS grad × age	1227***
	(12342)		(425)
		College grad × age	6109***
Marriage timing (base: after birth)			(717)
Marry before birth	39	Age^2	2
	(1341)		(8)
		HS grad × age^2	-8
Education × marriage timing			(6)
HS grad × marry before	14345***	College grad × age^2	-56***
	(1456)		(10)
College grad × marry before	19029***		
	(2569)	Constant	13745*
			(7673)
Observations			36,767
Adjusted R^2			0.159

Notes: Robust standard errors in parentheses; year fixed effects included. Dependent variable is husband's annual wage-and-salary income in 2016 dollars. The spouse's income, weeks and hours are reported for the calendar year before the interview; each observation is deflated with that year's CPI and dated by the wife's age in that year. Wife's education serves as proxy for husband characteristics through assortative mating. Baselines: high-school dropout (HSD) and first marriage after first birth. Coefficients and standard errors rounded to nearest dollar. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the survey population has not reached that age.

To ensure computational tractability, I model retirement income in reduced form as an education-specific replacement rate applied to pre-retirement earnings capacity, separately for women and husbands/partners.

Let T_R denote the first retirement period ($t = 14$; the last four model periods are retired). For each education group $e \in \{\text{HSD,HSG,COL}\}$, I compute a baseline pre-retirement earnings level as the average predicted annual labor income in the final working period,

$$\bar{w}_e \equiv \mathbb{E}[\widehat{w}_{it} \mid \text{Educ}_i = e,\ t = T_R - 1]\,, \qquad \bar{w}_e^m \equiv \mathbb{E}[\widehat{w}_{it}^m \mid \text{Educ}_i = e,\ t = T_R - 1]\,,$$

where the mean is an unweighted average over the model's wage grid in that period—every ability quartile, experience level and other discrete state relevant for the wage counts equally—rather than an average over the simulated population. The same base is used for every woman of a given education, whatever her marital and fertility history.

In retirement periods $t \geq T_R$, labor income is replaced by a deterministic benefit level:

$$w_e^R = \varrho_e\, \bar{w}_e, \qquad (w^m)_e^R = \varrho_e\, \bar{w}_e^m,$$

held constant over all retirement ages.

Social Security replacement rates. To discipline retirement income in the model, I calibrate education-specific replacement rates using Social Security Administration Office of the Chief Actuary replacement-rate statistics (first-year retired-worker benefits as a percent of wage-indexed career-average earnings). The model does not implement the statutory benefit formula (AIME/PIA) directly; instead, I use these statistics to discipline education-specific multipliers ϱ_e in a reduced-form retirement-income rule. In particular, ϱ_e should be interpreted as a replacement rate relative to late-career earnings in the model, proxied by $\bar{w}_e$, rather than literally relative to wage-indexed career-average earnings; the values used are $\varrho_e = 0.548$, 0.408 and 0.336 for dropouts, high-school graduates and college graduates.[1]

[1] Because $\bar{w}_e$ is a proxy for earnings capacity at the end of the working life rather than AIME, the mapping from the SSA tables into ϱ_e is an approximation that preserves the education gradient in replacement rates while maintaining a parsimonious retirement-income process.

S.5 The Pill-Share Re-estimation

The wantedness and within-method failure moments are plotted, model against data, in Figure 2 of the paper, and both models are set beside the data on them in Table OA.36. Tables S.18 and S.19 report the re-estimation with the pill shares given positive weight and its counterfactuals, discussed in Online Appendix OA.12.3.

Table S.18. What matching the pill share would cost: method costs re-estimated with the share targeted

	Estimate	Method costs refit to the pill share
Criterion (138 targeted moments)		
Total	458.5	542.3
First births by age	136.8	164.0
Unintended first birth by 21	4.7	22.8
Composition by age at first birth	29.9	42.6
Contraceptive use	84.2	89.5
Within-method failure	28.1	25.2
Labor	122.6	131.1
Schooling	37.8	53.0
Marriage	13.0	12.6
Child investment	1.4	1.5
Pill share among users, 18–25 (data 0.80/0.78/0.77/0.71, *Q1–Q4)*		
Q1 / Q2 / Q3 / Q4	0.50 / 0.50 / 0.56 / 0.57	0.71 / 0.72 / 0.76 / 0.76
Pill share among users, 26–33 (data 0.76/0.73/0.72/0.65*)*		
Q1 / Q2 / Q3 / Q4	0.43 / 0.45 / 0.49 / 0.51	0.66 / 0.68 / 0.70 / 0.72
Unintended first birth by 21, Q1–Q4 (data 0.33/0.25/0.17/0.10*)*		
Q1 / Q2 / Q3 / Q4	0.32 / 0.23 / 0.18 / 0.10	0.29 / 0.21 / 0.16 / 0.09
Method-cost parameters		
κ_{low}	0.0161	0.0522
κ_{pill}	0.0372	0.0270
σ_n (method taste shocks)	0.0806	0.0716
$\Delta\kappa_{pill}$ HS	0.0161	-0.0163
$\Delta\kappa_{pill}$ Coll	-0.0656	-0.0716

Notes: The first column is the estimate. The second re-optimizes the low-tier cost κ_1, the pill's cost κ_2 and its two education shifters, and the method taste-shock scale σ_n, with the eight pill-share moments given their inverse-variance weight and every other parameter held at its estimate; the criterion rows are evaluated on the 138 moments of the estimation, so they measure what the fit of the paper's targets gives up. The refit vector is a sensitivity, not an estimate.

Table S.19. Counterfactual experiments at the refit vector

	Baseline	Adherence = Q4	Cost of use: Matched	Cost of use: Partial cut	LARC menu: Colorado	LARC menu: Pill's cost	LARC menu: All users
Panel A: Aggregate							
First birth by 17	0.141	0.114	0.122	0.126	0.125	0.111	0.085
First birth by 21	0.364	0.315	0.315	0.327	0.339	0.306	0.262
First birth by 25	0.575	0.530	0.508	0.526	0.552	0.515	0.470
Unintended share	0.349	0.295	0.374	0.363	0.320	0.278	0.224
Any method, 18–25	0.741	0.746	0.807	0.791	0.747	0.756	0.757
Pill share among users	0.755	0.737	0.703	0.715	0.515	0.083	0.703
LARC share among users	0.000	0.000	0.000	0.000	0.249	0.702	0.000
College attendance	0.366	0.370	0.374	0.372	0.370	0.375	0.380
Employment, 26–45	0.531	0.531	0.535	0.533	0.531	0.531	0.531
Own earnings at 30	16.091	16.064	16.429	16.287	16.129	16.116	16.116
Panel B: First birth by 21, by AFQT quartile							
Q1	0.522	0.410	0.479	0.490	0.465	0.435	0.366
Q2	0.429	0.382	0.366	0.381	0.412	0.374	0.328
Q3	0.285	0.246	0.240	0.253	0.261	0.226	0.186
Q4	0.220	0.220	0.176	0.187	0.217	0.192	0.169
Cut in cost of use, δ	–	–	77.5%	59.0%	–	–	–

Notes: Table 6 recomputed at the vector of Table S.18, in the reported specification. Adherence, the failure technology and every parameter outside the method-cost block are the estimate's.

S.6 Counterfactuals with Unrestricted Teen Adherence and Realized Welfare

Tables S.20 and S.21 repeat the experiments of Section 6 of the paper with adherence at ages 14–17 left at its estimated values rather than held at each quartile's own 18–25 level (Section 6); the paper reports the restricted version and treats these as an upper bound. Table S.22 reports the realized-utility version of the welfare measure, which omits the taste shocks that the ex-ante measure of Section 6 integrates over.

Table S.20. Counterfactual experiments, unrestricted teen adherence (upper bound)

	Baseline	Adherence = Q4	Cost of use: No cost of use	Cost of use: Partial cut	LARC menu: Colorado	LARC menu: Pill's cost	LARC menu: All users
Panel A: Aggregate							
First birth by 17	0.161	0.082	0.139	0.144	0.130	0.121	0.073
First birth by 21	0.383	0.292	0.331	0.345	0.348	0.328	0.246
First birth by 25	0.589	0.509	0.520	0.538	0.561	0.537	0.447
Unintended share	0.406	0.312	0.425	0.420	0.366	0.342	0.248
Any method, 18–25	0.744	0.746	0.807	0.790	0.747	0.754	0.759
Pill share among users	0.557	0.532	0.619	0.603	0.321	0.057	0.492
LARC share among users	0.000	0.000	0.000	0.000	0.251	0.540	0.000
College attendance	0.364	0.378	0.381	0.376	0.370	0.374	0.383
Employment, 26–45	0.530	0.532	0.535	0.533	0.531	0.531	0.531
Own earnings at 30	16.099	16.121	16.538	16.401	16.135	16.141	16.092
Panel B: First birth by 21, by AFQT quartile							
Q1	0.581	0.378	0.527	0.538	0.501	0.480	0.340
Q2	0.455	0.357	0.398	0.414	0.420	0.396	0.307
Q3	0.296	0.232	0.244	0.259	0.271	0.247	0.174
Q4	0.202	0.202	0.159	0.169	0.200	0.189	0.161
Cut in cost of use, δ	–	–	100.0%	74.2%	–	–	–

Notes: As Table 6, but with adherence at ages 14–17 left at its estimated values rather than held at each quartile's own 18–25 level. The top quartile's teen adherence is 0.998 there, at a corner with a numerically zero weighted Jacobian column, and no moment in the estimation identifies it, so every experiment takes its largest single step on a cell the data do not discipline. These numbers are therefore an upper bound on what acting on adherence can achieve, and Table 6 is the specification the paper reports. The corresponding welfare numbers are in Table S.21.

Table S.21. Consumption-equivalent variation, unrestricted teen adherence (upper bound)

	Adherence = Q4	Cost of use: No cost of use	Cost of use: Partial cut	LARC menu: Colorado	LARC menu: Pill's cost	LARC menu: All users
Q1	57.4	42.5	27.3	18.0	28.2	76.3
Q2	32.3	61.6	37.8	7.3	15.7	48.8
Q3	29.6	126.1	67.5	5.2	16.8	63.9
Q4	0.0	146.8	75.4	0.2	4.9	16.3
All	28.4	80.9	47.6	7.9	16.5	49.6

Notes: As Table 7, with adherence at 14–17 left at its estimated values. The lowest quartile's gain under the targeted experiment is 57.4 percent here against 29.7 in the reported specification: bounding the one adherence cell no moment identifies halves it. The qualitative pattern is the same in both—targeted beats uniform at the bottom, and the uniform instrument's gains accrue to the middle and top. All three caveats in the text apply here too.

Table S.22. Consumption-equivalent variation on realized utility paths, by quartile and final education

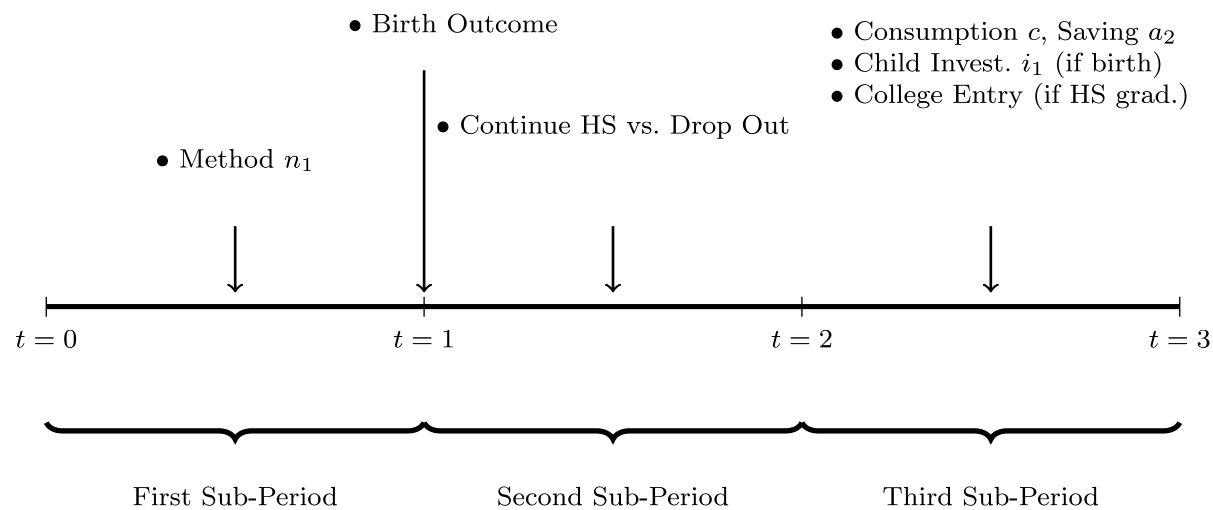
	Adherence = Q4	Cost of use: No cost of use	Cost of use: Partial cut	LARC menu: Colorado	LARC menu: Pill's cost	LARC menu: All users
By AFQT quartile						
Q1	-17.4	17.3	10.5	-13.5	-12.4	-24.6
Q2	-8.7	25.3	14.8	-8.2	-9.1	-17.0
Q3	-6.3	63.3	34.3	-10.2	-8.6	-14.7
Q4	0.0	53.8	29.3	-3.1	-6.2	-12.6
By final education						
HS dropout	-9.4	9.1	4.6	-7.3	-7.3	-16.2
High school	-9.8	30.7	18.2	-9.3	-9.8	-19.0
College	-5.1	125.1	57.0	-10.2	-8.9	-13.5
All	-9.0	36.0	20.7	-9.2	-9.3	-17.8

Notes: The realized-utility counterpart of Table 7, in the reported specification (teen adherence held at each quartile's 18–25 level): the permanent proportional change in consumption that equates realized lifetime utility under the baseline to realized lifetime utility under the experiment, computed path by path under common random numbers and averaged within group. Realized utility sums the utility of consumption, the net utility of work, the utility of a child at home, and the common cost of the method chosen each period with a negative sign. Education is final education, which the experiments move, so the rows are not fixed populations. It omits the pill's education shifter and the one-time utility terms attached to schooling choices, and it omits the taste shocks that drive method choice, so it understates the value of a new option to the women who choose it for its draw; for the two experiments that change no menu it is close to the ex-ante measure, and for the long-acting method it is below it.

S.7 Computational Appendix

This section documents the numerical solution, simulation and estimation procedures in full; Online Appendix OA.11.6 gives the summary. Figure S.2 shows the within-period timing of the teen period.

Figure S.2. Teenage Women (Ages 14–17): Within-Period Timing



Notes: Within-period timing at $t = 1$ (ages 14–17). All women are in high school. Sub-period (i): choose a contraceptive method n_1, then a stochastic birth may occur. Sub-period (ii): after observing the birth outcome, choose whether to continue high school or drop out; having a child raises the psychic cost of continuing (κ_{HS}). Sub-period (iii): consumption–saving, child investment if a birth occurred, and (for high-school graduates) the college-entry decision.

This appendix documents the numerical solution, simulation, and estimation procedures. The model is a finite-horizon discrete-time dynamic programming problem with both continuous (assets) and discrete (education, marital status, fertility) state variables. I solve the model by backward induction, using different numerical methods at different life stages depending on the structure of the within-period problem. The estimation uses the simulated method of moments (SMM), with a global optimization algorithm to search over the parameter space.

S.7.1 State space, grids, and timing

Time is discrete in four-year periods, indexed by $t = 1, \ldots, T$. The mapping from period to age is $\text{age}_t = 10 + 4t$, so $t = 1$ corresponds to age 14.

The individual state is

$$s_t \equiv (a_t, \theta, e_t, x_t, m_t, mk_t, k_t, t),$$

where a_t is assets at the beginning of t, θ is cognitive ability type (discrete), e_t is education (dropout / HS / college), x_t is experience (discrete, accumulated when working), m_t is marital status (single/married), mk_t is an indicator for whether the first birth occurred out of marriage, and k_t is child status. In the implementation, $k_t \in \{1, 2, 3\}$ corresponds to: no child; newborn in the current period; and older child in later periods.

The continuous state a_t is discretized on an exogenous grid $\mathcal{A} = \{a^1, \ldots, a^{N_a}\}$ with cubic spacing:

$$a_j = a_{\min} + (a_{\max} - a_{\min}) \cdot (j/N_a)^3, \quad j = 0, \ldots, N_a.$$

This concentrates grid points near the borrowing constraint where policy functions are steepest. Policy functions are stored on $\mathcal{A}$ and evaluated off-grid by linear interpolation in simulation.

Within-period timing and sub-stages. The code solves a three-substage problem within each fertile working period:

1. **Stage 3:** Given marital status and realized fertility outcome (child/no child), the household chooses labor $l_t \in \{0, 1\}$, savings a_{t+1}, consumption c_t, and (if a newborn arrives) child investment i_t.

2. **Stage 2:** Prior to the fertility realization, the household chooses a contraceptive method n_t, which governs the birth probability; the stage-2 value is the log-sum over methods of the stage-3 values integrated over the birth realization.

3. **Stage 1:** If single and eligible to meet, the household draws a meeting opportunity and chooses whether to marry; the stage-1 value integrates the stage-2 value over meeting opportunities and the marriage decision rule.

S.7.2 Household problem and key first-order conditions

Preferences are CRRA in consumption, $u(c_t) = c_t^{1-\gamma}/(1-\gamma)$, where γ is the coefficient of relative risk aversion for the woman. Per-adult-equivalent consumption is implemented via an equivalence-scale denominator

$$\phi_c(m_t, b) = 1 + \mathbf{1}\{m_t = \text{married}\}\phi_{ca} + b\,\phi_{ck}, \quad b = \mathbf{1}\{k_t = 2\}.$$

Thus, the child-related equivalence-scale term ϕ_{ck} enters the budget constraint only in the birth period ($k_t = 2$), consistent with the "one-period child in the household" assumption. After the birth period, the state moves from $k_t = 2$ to $k_{t+1} = 3$ (child has left the household), so that $\mathbf{1}\{k_{t+1} = 2\} = 0$ in all subsequent periods.

Let y_t denote disposable (post-tax/post-transfer) income in period t (four-year total). Gross income is transformed by a progressive tax-transfer function:

$$y_t = \lambda_\tau \cdot \text{gross}_t^{1-\tau} + \underline{y}(m_t),$$

where $\tau = 0.18$ is the progressivity parameter, $\lambda_\tau = 0.85$ is the scale parameter, and $\underline{y}(m_t)$ is the guaranteed minimum income (\$8,634 for singles, \$12,943 for couples, yearly in 2016 dollars).

The stage-3 budget constraint is

$$c_t \cdot \phi_c(m_t, b) + i_t + a_{t+1} = (1+r)a_t + y_t.$$

Child investment subproblem (stage 3, newborn only). When $k_t = 2$ (newborn in period t), child investment enters the continuation value through

$$V_k(i_t) = \omega_{0,\theta} + \omega_1 i_t^{\omega_2}, \quad \omega_2 < 1.$$

The household's problem is

$$\max_{c_t, i_t} \quad u(c_t) + V_k(i_t) + \beta V_{t+1}(a_{t+1}) \quad \text{s.t.} \quad c_t \cdot \phi_c(m_t, b) + i_t + a_{t+1} = (1+r)a_t + y_t.$$

The first-order condition equates marginal utility per dollar:

$$\frac{u'(c_t)}{\phi_c(m_t, b)} = V_k'(i_t) \quad \Longleftrightarrow \quad c_t^{-\gamma} = \phi_c(m_t, b)\,\omega_1\omega_2\, i_t^{\omega_2-1}.$$

Solution method. Given a_{t+1}, the budget constraint implies $c_t \cdot \phi_c + i_t =$ available, where available $\equiv (1+r)a_t + y_t - a_{t+1}$. Substituting into the FOC yields a single equation in i_t. The code solves this via bisection on $i_t \in [10^{-6}, 0.9999 \times \text{available}]$:

1. Compute $c_t(i_t) = (\text{available} - i_t)/\phi_c$.
2. Evaluate FOC residual: $r(i_t) = c_t(i_t)^{-\gamma} - \phi_c(m_t, b)\,\omega_1\omega_2\, i_t^{\omega_2-1}$.
3. Update bracket: if $r(i_t) < 0$, increase i_t (consumption too high); else decrease.
4. Terminate when $|r(i_t)| < 10^{-10}$ or bracket width $< 10^{-10}$ (max 50 iterations).

The solution is unique because $r(i_t)$ is strictly increasing in i_t: $c_t(i_t)$ is decreasing in i_t, so $u'(c_t(i_t))$ rises, while $V_k'(i_t)$ falls when $\omega_2 < 1$.

Contraception choice (stage 2). Given stage-3 values with and without a birth, $(V_t^{3,j=2}, V_t^{3,j=1})$, and ΔV_t their difference, the stage-2 value is the log-sum over the three methods of Section 3.4,

$$V_t^2 = V_t^{3,j=1} + \mathbb{E}_{\varepsilon^\omega}\left(\sigma_n \log \sum_{n=0}^{2} \exp\left\{\frac{p_n\,\tilde{\Delta V}_t - \kappa_n}{\sigma_n}\right\}\right),$$

where $\tilde{\Delta V}_t = \Delta V_t + \sigma_\omega \varepsilon_{it}^\omega$ and the expectation is over the transitory wantedness shock (seven-point Gauss–Hermite), mixed with the exposure share $\pi_{\theta t}$ as in Online Appendix OA.7.5. The choice probabilities are the logit shares evaluated at the realized $\tilde{\Delta V}_t$; the simulation draws the method from them and then the birth from that method's probability p_n (the expected probability $\bar{p}_t = \sum_n P(n) p_n$ is what the woman faces before the draw); no first-order condition is involved.

Labor choice with taste shocks. In periods solved by DC-EGM, the labor decision has i.i.d. Type-I extreme value taste shocks with scale σ_l, implying an inclusive value (log-sum) aggregator and a logit work probability:

$$V_t = \sigma_l \log\Big(\exp(V_{t,l=0}/\sigma_l) + \exp(V_{t,l=1}/\sigma_l)\Big),$$
$$P_t(l=1) = \frac{\exp(V_{t,1}/\sigma_l)}{\exp(V_{t,0}/\sigma_l) + \exp(V_{t,1}/\sigma_l)}.$$

S.7.3 Solution algorithm (backward induction)

This section documents the solver `VFI_P_DCEGM` in `vfi_dcegm.jl`. The algorithm proceeds by backward induction, but uses different numerical routines depending on age.

Overview. Let $N_R = 4$ be the number of retired periods, and let $N_{NF} = 6$ denote the number of working periods after fertility ends. The code partitions the horizon into: (i) retirement ($t > T - N_R$), solved by EGM; (ii) non-fertile working ages ($T - N_R - N_{NF} < t \leq T - N_R$), solved by DC-EGM; (iii) fertile ages ($t \leq T - N_R - N_{NF} = 7$), solved by VFI with grid search (plus a one-dimensional inner problem for i_t and the closed-form method log-sum).

The rationale for this partition is computational efficiency. In retirement, there is no labor-leisure choice and the problem reduces to a standard consumption-saving model, for which EGM is highly efficient. In non-fertile working ages, the discrete labor choice introduces non-convexities that can generate discontinuous policy functions; DC-EGM handles this by computing choice-specific consumption functions via EGM and then applying an upper-envelope algorithm to recover the global optimum. In fertile ages, the additional discrete margins (marriage, contraception, schooling) and the multi-stage within-period structure make DC-EGM less practical, so the code reverts to grid search with analytical solutions for the continuous sub-problems.

Algorithm 1 (Retirement, EGM). In retirement, labor is absent and the problem is a standard consumption-saving model with a borrowing constraint. The EGM step for each discrete state (θ, e, m, mk, k) is:

1. Fix the exogenous grid for next-period assets $\mathcal{A} = \{a'\}$.

2. For each $a' \in \mathcal{A}$, compute expected marginal utility next period using the already-solved consumption policy $c_{t+1}(\cdot)$, and invert the Euler equation

$$u'(c_t(a')) = \beta(1+r)\,\mathbb{E}[u'(c_{t+1}(a'))]$$

to obtain $c_t(a')$.

3. Use the budget constraint to map $(a', c_t(a'))$ into the endogenous current asset level $a_t(a')$.

4. Interpolate from the endogenous grid back to the exogenous grid, impose the borrowing constraint, and store $c_t(a)$, $a_{t+1}(a)$, and $V_t(a)$.

Algorithm 2 (Non-fertile working ages, DC-EGM). In working ages after fertility ends ($t \in \{T - N_R - N_{NF} + 1, \ldots, T - N_R\}$), the household chooses labor $l_t \in \{0, 1\}$ and savings. Because labor is discrete and shocks are extreme value, the continuation value involves an inclusive value and choice probabilities. The code implements DC-EGM following Iskhakov et al. (2017).

For each period t (going backward) and each discrete state (θ, e, x, m, mk, k):

1. *Choice-specific EGM step.* For each current labor choice $l_t \in \{0, 1\}$:

   (a) Compute disposable model-period income $y_t(l_t) = 4\big[\lambda_\tau \big(\tilde{y}_t^0(l_t)\big)^{1-\tau} + \underline{y}(m_t)\big]$ from gross annual household income $\tilde{y}_t^0(l_t)$, with the progressive tax-and-transfer schedule of the main text.

   (b) For each $a' \in \mathcal{A}$ (exogenous next-period asset grid), compute expected marginal utility at $t+1$:

   $$\mathbb{E}_t[u'(c_{t+1})] = \sum_{l'=0}^{1} P_{t+1}(l' \mid a') \cdot u'(c_{t+1,l'}(a')),$$

   where $P_{t+1}(l' \mid a')$ is the labor-choice probability from the previous iteration (logit), so that $P_{t+1}(1 \mid a')$ is the work probability and $P_{t+1}(0 \mid a') = 1 - P_{t+1}(1 \mid a')$.

(c) Invert the Euler equation to obtain consumption on the endogenous grid:

$$c_{t,l_t}(a') = \left[\beta(1+r)\,\mathbb{E}_t[u'(c_{t+1})]\right]^{-1/\gamma}.$$

(d) Map to endogenous current assets using the budget constraint:

$$a_{t,l_t}(a') = \frac{c_{t,l_t}(a') \cdot \phi_c(m,b) + a' - y_t(l_t)}{1+r}.$$

(e) Construct the choice-specific value on the endogenous grid:

$$V_{t,l_t}\big(a_{t,l_t}(a')\big) = u(c_{t,l_t}(a')) + \mathbf{1}\{l_t = 1\}\psi_l(t,e) + \beta\, V_{t+1}(a').$$

2. *Upper envelope.* The endogenous grid $(a_{t,l}, c_{t,l}, V_{t,l})$ may be non-monotonic when labor decisions change discontinuously. Apply the upper-envelope method:

(a) Sort by endogenous assets $a_{t,l}$.

(b) Check monotonicity: if $a_{t,l,j+1} \geq a_{t,l,j} - 10^{-10}$ for all j, use direct interpolation.

(c) Otherwise, for each exogenous grid point $a \in \mathcal{A}$, compute $V_t(a) = \max_j V_{t,l}(\text{segment}_j(a))$ over all segments.

3. *Credit constraint region.* For $a < \min(\{a_{t,l}(a')\})$, set $c_t = (a(1+r) + y_t - \underline{a})/\phi_c$ and $a_{t+1} = \underline{a}$.

4. *Logit aggregation.* Aggregate choice-specific values with Type-I EV taste shocks (scale σ_l):

$$V_t(a) = \sigma_l \log\big(\exp(V_{t,0}(a)/\sigma_l) + \exp(V_{t,1}(a)/\sigma_l)\big),$$
$$P_t(l = 1 \mid a) = \frac{\exp(V_{t,1}(a)/\sigma_l)}{\exp(V_{t,0}(a)/\sigma_l) + \exp(V_{t,1}(a)/\sigma_l)}.$$

Algorithm 3 (Fertile ages and schooling, VFI with grid search). In fertile ages (and in early schooling periods), the code switches to grid-search VFI because the within-period structure (meeting/marriage, contraception and birth risk, newborn investment, schooling decisions, and experience dynamics) generates non-convexities and additional discrete margins that are not well suited for DC-EGM.

For each fertile period t (going backward) and each discrete state (θ, e, x, m, mk, k):

1. *Stage 3 (given marital and fertility outcome).* For each labor choice $l_t \in \{0, 1\}$, the code searches over $a_{t+1} \in \mathcal{A}$ and computes implied consumption from the budget. If $k_t = 2$ (newborn), it solves (c_t, i_t) jointly using the FOC (bisection method described above) for each candidate a_{t+1}. It stores the maximizing a_{t+1}, c_t, i_t and the resulting choice-specific value.

2. *Labor aggregation.* For each state, it aggregates across l_t using the log-sum formula with scale σ_l.

3. *Stage 2 (contraception and birth risk).* For states with no child ($k_t = 1$), it computes $V_t^{3,j=2}$ and $V_t^{3,j=1}$ from stage 3 and forms the method log-sum in closed form. It then forms the expected value integrating over the realized birth.

4. *Stage 1 (meeting and marriage).* For eligible singles, it applies the meeting probability $\mu_{t,e}$ and compares the stage-2 value under marriage versus remaining single, generating the marriage policy and the beginning-of-period value.

5. *Schooling decisions.* In the first periods, it solves high-school continuation and college attendance/continuation decisions using choice-specific value comparisons with extreme-value taste shocks.

Numerical details. (i) Grid search is accelerated by breaking when consumption turns negative and by exploiting local monotonicity in a'. (ii) The child-investment inner problem uses bisection with tolerance 10^{-10} and maximum 50 iterations. (iii) All consumption values are floored at 10^{-10} before utility evaluation to prevent numerical overflow.

Convergence and numerical tolerances. The solver employs the following numerical tolerances:

- *Consumption positivity:* $c_t \geq 10^{-10}$ (machine epsilon floor)

- *Child investment FOC:* Bisection terminates when $|u'(c_t) - V_k'(i_t)| < 10^{-10}$ or bracket width $< 10^{-10}$ (maximum 50 iterations)
- *Upper envelope:* Segments are considered monotonic if $a_{t,j+1} - a_{t,j} > -10^{-10}$
- *Interpolation:* Weights clamped to $[0,1]$ using $w = \min(\max(w,0),1)$
- *Logit aggregation:* Uses log-sum-exp trick to prevent overflow: $\log(\sum_j \exp(x_j)) = x_{\max} + \log(\sum_j \exp(x_j - x_{\max}))$

S.7.4 Forward simulation

The function `simulationF` takes the policy objects produced by `VFI_P_DCEGM` and simulates N life histories. It uses pre-drawn uniform random variables for fertility, meeting, and labor choices to ensure reproducibility across parameter vectors. In early periods, schooling continuation and college continuation/dropout are stage-3 policies that are indexed by the realized fertility outcome $j \in \{k, nk\}$; accordingly, these schooling rules are evaluated after the fertility draw and conditional on the realized j (see Section 3 and Appendix S.7.3).

Algorithm 4 (Simulation). For each simulated woman $i = 1, \ldots, N$:

1. Initialize $(a_1, \theta, e_1, x_1, m_1, mk_1, k_1)$ and store deterministic objects (age mapping, IDs).
2. For $t = 1, \ldots, T$:
   (a) Evaluate policy functions at the current asset level by linear interpolation on $\mathcal{A}$.
   (b) If eligible and single, realize a meeting draw and apply the marriage decision rule (sub-stage 1).
   (c) If in fertile ages and without a child, draw the method from its logit probabilities, compute p_n, and realize conception with the fertility draw (sub-stage 2), obtaining $j \in \{k, nk\}$.
   (d) Apply schooling decisions in early periods using the stage-3 policy rules conditional on the realized j (high-school continuation, college attendance/continuation/dropout).

(e) Realize labor supply using $P_t(l = 1)$ and the labor draw. Update experience deterministically when working.

(f) Given realized discrete outcomes, update assets using the savings policy; store consumption, income, and other outcomes.

3. After simulating all individuals, compute model moments from simulated histories.

S.7.5 Calibration (SMM) and optimization

I estimate the model using the simulated method of moments (SMM). SMM chooses parameters to minimize the distance between moments computed from simulated model output and their empirical counterparts.

Target moments and loss function. The Stata pipeline exports 153 targets: 10 schooling and early-fertility moments, 2 child-investment ratios, 7 Q1/Q4 birth-rate ratios, 28 first-birth shares by age and ability quartile, 17 marriage shares, 36 employment rates, 25 contraception-use rates by education, age and ability half, 4 unintended-first-birth shares by quartile, 16 method-choice moments (within-method failure and pill shares by quartile), and 8 unintended shares of first births by age at first birth for the lowest and highest quartiles. The seven ratios are exact functions of the quartile birth shares already in the vector and are dropped at load, and the eight pill-share moments are carried with zero weight (Section 4), so the estimator sees $J = 146$ moments of which 138 carry positive weight. Let $m^{\text{data}} \in \mathbb{R}^{146}$ denote the empirical moments and $m(\Theta)$ the simulated moments under parameter vector Θ. The SMM criterion is

$$\mathcal{L}(\Theta) = \sum_{j=1}^{146} w_j \big(m_j(\Theta) - m_j^{\text{data}}\big)^2, \qquad w_j = \frac{1}{\hat{\sigma}_j^2},$$

where $\hat{\sigma}_j$ is the bootstrap standard error of data moment j from 1,000 replications that resample women and rebuild every moment on the same draw, so that the bootstrap also delivers the full covariance matrix $\hat{\mathcal{V}}$ of the moment vector. The weights are fixed: they depend on the data alone and not on Θ, which is what the standard SMM asymptotics require (McFadden, 1989; Pakes and Pollard, 1989) and what makes the sandwich variance well defined. Two choices call for comment. First, I use the diagonal of $\hat{\mathcal{V}}^{-1}$ rather than the full matrix. $\hat{\mathcal{V}}$ is numerically singular (rank 136 of 138, condition number of order 10^{17}), so the optimal weighting matrix is not available, and Altonji and Segal (1996) show that the diagonal choice is preferable in samples of this size in any case. Second, no block receives a multiplier: each moment enters with the precision the data give it. This is the rule of the literature, and it is the only one that does not require a judgment about which block the estimator should care about most; the one cost is that the criterion is dominated by the blocks with the most precisely estimated cells, and I report the share of the criterion carried by each block alongside the fit. The two child-investment ratios are external constants from Caucutt and Lochner (2020), so the bootstrap returns no sampling variation for them and they are scaled by their empirical magnitude instead. Standard errors of the parameters are computed from the sandwich formula with the numerical Jacobian of $m(\Theta)$; the participation, pill-share and within-method-failure moments are averages of predicted probabilities across simulated women, which makes them smooth in Θ, while the birth, marriage and employment moments are averages of realized draws under common random numbers.

Choice of moments. The 146 moments the code carries—of which 138 are targeted—are the ten blocks listed above: schooling and early fertility (10), child investment (2), first-birth shares by age and ability quartile (28), marriage by age and education (17), employment by age and education (36), contraception use by age, education and ability half (25), unintended first-birth shares by quartile (4), unintended shares of first births by age at first birth (8), within-method failure by quartile (8), and the untargeted pill shares (8). Wage profiles are *not* targeted: the earnings process is estimated externally, outside the structural SMM step (Section 4.1).

Algorithm 5 (SMM objective evaluation). Given a candidate parameter vector Θ:

1. Map Θ into model objects (e.g., the conception technology parameters, labor preference/taste-shock scales, meeting probabilities, and child-investment parameters).

2. Solve the model to obtain value and policy functions (Algorithm 1–3).

3. Simulate outcomes (Algorithm 4).

4. Compute $m(\Theta)$ from simulated histories and return $\mathcal{L}(\Theta)$.

Global optimization and parallelization. The file `drivers/calibrate.jl` runs a global search with a self-adaptive differential evolution (jDE, rand/1/bin with self-adapting F and CR; Brest et al., 2006) in which each generation of trial vectors is evaluated in parallel across the workers of one node. Differential evolution is a derivative-free global optimizer well suited to non-convex, high-dimensional problems where the objective function is noisy or discontinuous. It maintains a population of candidate solutions and iteratively improves them through mutation, crossover, and selection operations. The algorithm operates as follows:

1. Initialize a population of $2\times$(number of workers) vectors around the warm start (three quarters as local perturbations, one quarter drawn uniformly on the bounds); 47 workers on a 48-core node give a population of 94.

2. Each generation, evaluate every trial vector in parallel—one solve–simulate–score per worker, about four seconds each on a single-threaded worker—keep the improvements, and write the incumbent and the population to disk every ten minutes.

3. Terminate at the wall-clock budget, or when the population has collapsed on the incumbent (median within 0.2% of the best). The production runs were self-limited to 36 hours for the unrestricted model and 28 for the restricted one, and each was then re-estimated for 24 hours from its incumbent on the final moment vector; over the last six hours of that polish the criterion moved by less than half a percent, which is the convergence check I report alongside the multi-start evidence.

S.7.6 Computational performance and implementation

Hardware and software. Estimation was performed on a high-performance computing cluster with Intel Xeon Gold 6248R processors (48 cores per node, 3.0 GHz base frequency).

The code is implemented in Julia 1.11 (cluster) and 1.12 (laptop), with a pinned package manifest in the replication package, leveraging multithreading for the EGM/DC-EGM steps and `Distributed.jl` for the calibration.

Solution time. One complete evaluation at the estimated parameters—backward induction on the 30-point asset grid, simulation of 10,000 women, and construction of the 146 moments—takes about 1.2 seconds on four threads and about four seconds on a single thread. The estimation runs single-threaded workers, so the second figure is the one that sets the pace of the search; the first is what a replicator sees on a laptop.

Calibration runtime. The SMM estimation uses the parallel self-adaptive differential evolution described above, with 47 workers and a population of 94 on one 48-core node, sustaining about 12.0 evaluations per second. The unrestricted estimate ran for 36 hours (1,539,814 evaluations) and the restricted one, started from the best of ten random starts, for 28 hours (1,192,014); each was then re-estimated for 24 hours from its incumbent on the moment vector of the unified sample (1,032,496 and 1,020,088 evaluations), a polish in which the criterion moved by less than half a percent over the last six hours. The random and near starts of Online Appendix OA.11.4 consumed about twice that budget again.

Grid density and accuracy. The baseline specification uses $N_a = 30$ asset grid points with cubic spacing, $a_j \propto j^3$, to concentrate points near the borrowing constraint; the child-investment and contraception grids are also 30 points. The grid can be refined uniformly from the command line, and `drivers/jacobian_step_scan.jl` was written to test whether the non-smoothness that complicates the numerical Jacobian originates in the solution rather than in the simulation: raising the number of simulated women twentyfold did not make the finite differences converge, which rules out sampling, while the policy is interpolated on a finite grid and the discrete-continuous upper envelope compares segments, so optimal choices move in jumps as parameters vary for any number of agents. This is why the standard errors are computed on 160,000 simulated women and reported at two finite-difference steps. Child investment is solved analytically via the first-order condition (bisection with tolerance 10^{-10}), avoiding discretization error.

Numerical stability. To ensure stability: (i) All consumption values are floored at 10^{-10} before utility evaluation. (ii) Logit aggregation uses the log-sum-exp trick: $\log(\sum_j \exp(x_j)) = x_{\max} + \log(\sum_j \exp(x_j - x_{\max}))$ to prevent overflow. (iii) Interpolation weights are clamped to $[0, 1]$. (iv) The bisection algorithm for child investment uses robust bracketing with explicit checks for corner solutions.

Computational requirements.

- *Minimal replication:* one solve, simulate and score at the estimated vector takes a few seconds on a standard laptop (4 cores, 16GB RAM); `main.jl` reproduces every table and figure in Section 5 from `best_sol.txt` in under a minute.
- *Full estimation:* one 48-core node. The production pair was requested with 180GB and a seven-day limit and used eight hours of wall clock per variant, in two jobs.
- *Post-estimation:* the numerical Jacobians for both variants, at two finite-difference steps and 160,000 simulated women, plus the nested test, take 16 minutes on 24 threads and 64GB.

Random number generation. All stochastic elements (simulation draws for fertility, marriage, labor, education) use pre-generated uniform random variables with fixed seed (4546), ensuring exact replicability across parameter vectors. This design ensures that changes in moments reflect only parameter changes, not simulation noise. Calibration uses pseudo-random perturbations for initial parameter values (seed set per worker ID).

Software dependencies. Core packages with versions: `Parameters.jl` (0.12), `Interpolations.jl` (0.14), `BlackBoxOptim.jl` (0.6), `Distributed.jl` (standard library), `DataFrames.jl` (1.5), `Distributions.jl` (0.25), `CSV.jl` (0.10). Full environment specified in `Project.toml` in the replication package.

References

Joseph G. Altonji and Lewis M. Segal. Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics*, 14(3):353–366, 1996.

Janez Brest, Saso Greiner, Borko Bošković, Marjan Mernik, and Viljem Žumer. Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. *IEEE Transactions on Evolutionary Computation*, 10(6):646–657, 2006.

Pedro Carneiro and James J. Heckman. Human capital policy. In James J. Heckman and Alan B. Krueger, editors, *Inequality in America: What Role for Human Capital Policies?*, chapter 2, pages 77–239. MIT Press, Cambridge, MA, 2003.

Elizabeth M Caucutt and Lance Lochner. Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy*, 128(3):1065–1147, 2020.

Jonah B. Gelbach. When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, 34(2):509–543, 2016.

James J. Heckman, Jora Stixrud, and Sergio Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24 (3):411–482, 2006.

Fedor Iskhakov, Thomas H Jørgensen, John Rust, and Bertel Schjerning. The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics*, 8(2):317–365, 2017.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989. doi: 10.2307/1913621.

Derek A. Neal and William R. Johnson. The role of premarket factors in black–white wage differences. *Journal of Political Economy*, 104(5):869–895, 1996. doi: 10.1086/262045.

Ariel Pakes and David Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057, 1989. doi: 10.2307/1913622.

Michael S. Rendall, Lynda Clarke, H. Elizabeth Peters, Nalini Ranjit, and Georgia Verropoulou. Incomplete reporting of men's fertility in the United States and Britain: A research note. *Demography*, 36(1):135–144, 1999.

James Trussell. Contraceptive failure in the United States. *Contraception*, 83(5):397–404, 2011.