Revealed Beliefs and the Marriage Market Return to Education

Online Appendix

Alison Andrew & Abi Adams

A Additional Figures & Tables

Table A.1: Balance by whether respondent was assigned to ex-post vs. ex-ante survey instrument

	(1) Ex-post		(2) Ex-ante	
	Mean	SD	Mean	SD
Age in years	42.11	(8.306)	41.71	(8.424)
Own age at marriage in years*	15.50	(3.231)	15.64	(3.486)
Years of school*	1.473	(3.260)	1.516	(3.277)
Can read complete sentence (in Hindi)*	0.0960	(0.295)	0.112	(0.316)
Number of sons*	2.121	(1.118)	2.112	(1.107)
Number of daughters*	2.441	(1.307)	2.452	(1.335)
Owns asset that can dispose of at will	0.136	(0.342)	0.129	(0.335)
Can go to market unaccompanied*	0.601	(0.490)	0.621	(0.485)
At least some say over when child gets married	0.964	(0.186)	0.961	(0.193)
At least some say over to whom child gets married	0.955	(0.207)	0.949	(0.219)
At least some say over when child leaves school	0.944	(0.230)	0.939	(0.239)
Has done any work (inc. on family farm) in last year	0.610	(0.488)	0.579	(0.494)
Has worked for cash in last year	0.355	(0.478)	0.334	(0.472)
Has child (male or female) who is married	0.369	(0.483)	0.357	(0.479)
House has dirt floor*	0.522	(0.500)	0.492	(0.500)
Scheduled caste or scheduled tribe*	0.362	(0.481)	0.342	(0.475)
Other Backward Caste or Economically Backward Class*	0.447	(0.497)	0.454	(0.498)
Hindu*	0.966	(0.180)	0.969	(0.172)
N	2324		2258	

Notes: Table reports descriptive statistics for: (1) caregivers assigned to ex-post survey instrument; (2) caregivers assigned to ex-ante survey instrument. *Variable measured in baseline survey during 2016. All other variables collected in 2017/18 survey.

Table A.2: Summary of Vignette Characteristics

Characteristic	**Support**
Characteristics of the Bride and Bride's Family	
Wealth	{Very Poor, Average, Very Wealthy}
Age of Daughter	Ex-Post: Always 12
	Ex-Ante: {13, 14, ..., 21, 22}
Currently in School	{In School, Out of School}
Likes School	{Like, Dislike}
School Costs	{Free, Costly}
Chores	{Relax, Help Needed}
Comply with Norms (Ex-ante only)	{Polite and Well Behaved, Friends with Boys}
Age and Education at Marriage	
Age at Marriage (girl)	{13, 14, ..., 21, 22}
School Grade at Marriage	{7, ..., 12, College}
Characteristics of the Marriage Match	
Wealth	{ Very Poor, Average, Very Wealthy}
Age at Marriage	{21, 22, ..., 29, 30}
School Grade at Marriage	{None, 1, 5, 7, ..., 12, College}
Occupation	{Government Job, No Gov Job}
Minimum Dowry Acceptable to Groom (lakh)	{0, 0.5, 1, ..., 7, 7.5}

Notes: Table gives the support of the different characteristics described in the hypothetical vignettes.

Table A.3: Robustness of Preference Estimates

	(1)		(2)	
	Baseline		**No inattention**	
Groom characteristics				
Groom has government job	1.000		1.000	
Groom's age	-0.035	(0.014)	-0.038	(0.015)
Groom's education	0.060	(0.016)	0.062	(0.018)
Groom college education	-0.021	(0.148)	0.032	(0.153)
Low wealth	0.000		0.000	
Medium wealth	0.232	(0.117)	0.223	(0.121)
High wealth	0.193	(0.137)	0.169	(0.141)
Dowry (lakh)	-0.052	(0.032)	-0.059	(0.033)
Age of Marriage				
Age	0.563	(0.076)	0.617	(0.089)
(Age-18) $\times$ 1(Age $\geq$ 18)	-0.514	(0.086)	-0.556	(0.097)
Education at Marriage				
College	0.168	(0.182)	0.138	(0.188)
Years of education - 7	0.105	(0.034)	0.135	(0.038)
Preference shifters				
'Daughter likes school' x Years of schooling	0.073	(0.037)	0.074	(0.037)
'School is costly' x Years of schooling				
'Help at home needed' x Years before marriage				
Fixed parameters				
σ_ν	1.202	(0.231)	2.662	(0.360)
Inattention share	0.367	(0.046)	0.000	
Number of experiments	6972		6972	
Number of respondents	2324		2324	

Notes: Column (1) reprints our main preference results (from Table II) while Column (2) shows equivalent estimates restricting the inattention parameter to be 0.

Table A.4: Robustness of Belief Estimates

	Baseline		**No inattention**		**Discount factor =0.99**		**Exclude Norms**		**Add $ed.^2$**		**No weight on $u(T)$ in $EV(T+1,.)$**		**Estimate weight on $u(T)$ in $EV(T+1,.)$**		**Offer prob. vary with age**	
	(1)		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**		**(7)**		**(8)**	
Age	-0.001	(0.010)	-0.066	(0.044)	-0.006	(0.011)	-0.005	(0.017)	0.032	(0.023)	0.000	(0.010)	-0.016	(0.015)	-0.009	(0.038)
Education	1.243	(0.175)	1.146	(0.154)	1.236	(0.143)	1.239	(0.159)	0.630	(0.348)	1.421	(0.240)	1.117	(0.204)	0.943	(0.490)
College	3.563	(0.981)	2.533	(0.313)	3.575	(0.736)	3.758	(0.909)	1.799	(1.350)	3.602	(0.615)	3.823	(0.955)	2.359	(1.189)
Age*Education	-0.100	(0.026)	-0.078	(0.018)	-0.099	(0.020)	-0.103	(0.021)	-0.120	(0.020)	-0.125	(0.030)	-0.089	(0.026)	-0.124	(0.030)
Constant	-3.398	(0.100)	-2.485	(0.434)	-3.363	(0.112)	-3.387	(0.164)	-3.222	(0.191)	-3.443	(0.104)	-3.266	(0.150)	-3.129	(0.334)
Education^2									0.163	(0.070)						
Offer Probability																
Base Probability (γ_0)	0.375	(0.044)	0.167	(0.015)	0.385	(0.034)	0.352	(0.035)	0.361	(0.034)	0.412	(0.024)	0.395	(0.042)	0.341	(0.042)
Breaking norms (γ_1)	-0.163	(0.069)	-0.178	(0.059)	-0.160	(0.050)	0.000		-0.177	(0.053)	-0.143	(0.059)	-0.160	(0.059)	-0.176	(0.056)
Age 18+ (γ_2)															0.105	(0.077)
$\boldsymbol{EV(T+1,\bar{\omega}_{T+1})}$																
ϕ	-5.926	(1.466)	-3.786	(0.553)	-6.670	(1.317)	-5.820	(1.598)	-6.019	(1.365)	2.490	(1.664)	-6.335	(1.305)	-6.399	(2.387)
Weight on $u(T,.,q=0)$	1.000		1.000		1.000		1.000		1.000		0.000		0.978	(0.088)	1.000	
Preference shocks																
σ_ε	0.600	(0.092)	3.082	(0.486)	0.522	(0.091)	0.605	(0.114)	0.666	(0.094)	0.628	(0.096)	0.589	(0.097)	0.659	(0.085)
Inattention share	0.367		0.000		0.367		0.367		0.367		0.367		0.367		0.367	
Number of experiments	5524		5524		5524		5524		5524		5524		5524		5524	
Number of respondents	2258		2258		2258		2258		2258		2258		2258		2258	

Notes: Column (1) reprints our main belief results (from Table III) while Columns (2) - (6) show equivalent estimates after altering different assumptions. In particular, in (2) we restrict the inattention parameter to 0, in (3) we use a discount factor of 0.99 rather than 0.95, in (4) we don't allow the offer probability to depend on whether a girl conforms to gender norms, in (5) we add a second order polynomial in education, in (6) we assume terminal values are independent of flow-payoffs in period T, in (7) we estimate the weight on this same flow payoff, and in (8) we allow the offer rate to depend on whether or not the daughter is aged 18+ (In particular, we set $1-\pi^N(\bar{\omega}) = \gamma_0(1+\gamma_1 Norms)(1+\gamma_2 \mathbf{1}(Age > 17))$.)

Table A.5: Determinants of Groom Choices: Reduced Form Probit

	(1)		(2)		(3)	
Bride's Age at Marriage=14	0.1021	(0.0765)	0.1173	(0.0769)	0.1429*	(0.0866)
Bride's Age at Marriage=15	0.2030***	(0.0710)	0.2191***	(0.0715)	0.2410***	(0.0813)
Bride's Age at Marriage=16	0.2563***	(0.0717)	0.2755***	(0.0722)	0.3329***	(0.0815)
Bride's Age at Marriage=17	0.2581***	(0.0716)	0.2678***	(0.0721)	0.3117***	(0.0810)
Bride's Age at Marriage=18	0.5207***	(0.0700)	0.5345***	(0.0704)	0.5976***	(0.0793)
Bride's Age at Marriage=19	0.5871***	(0.0750)	0.6062***	(0.0753)	0.6620***	(0.0851)
Bride's Age at Marriage=20	0.5850***	(0.0726)	0.6042***	(0.0731)	0.6790***	(0.0823)
Bride's Age at Marriage=21	0.5789***	(0.0697)	0.5933***	(0.0702)	0.6685***	(0.0793)
Bride's Age at Marriage=22	0.6506***	(0.0754)	0.6625***	(0.0759)	0.7612***	(0.0863)
Bride's Education =8	0.0854*	(0.0491)	0.0917*	(0.0499)	0.1098*	(0.0572)
Bride's Education =9	0.1023**	(0.0501)	0.1165**	(0.0521)	0.1028*	(0.0591)
Bride's Education =10	0.2611***	(0.0579)	0.2829***	(0.0607)	0.2435***	(0.0702)
Bride's Education =11	0.2456***	(0.0556)	0.2715***	(0.0596)	0.2261***	(0.0686)
Bride's Education =12	0.2876***	(0.0622)	0.3168***	(0.0678)	0.3009***	(0.0776)
Bride's Education =13	0.3117***	(0.0514)	0.3532***	(0.0621)	0.2731***	(0.0701)
Bride has Male Friends	-0.3283***	(0.0306)	-0.3248***	(0.0307)	-0.3343***	(0.0344)
Dowry	0.0598***	(0.0103)	0.0591***	(0.0103)	0.0552***	(0.0117)
Bride's Wealth = Poor	0.0071	(0.0374)				
Bride's Wealth = Wealthy	-0.0192	(0.0365)				
Bride's Wealth = Poor, Groom's Wealth = Poor			-0.0361	(0.0634)	0.0177	(0.0716)
Bride's Wealth = Wealthy, Groom's Wealth = Poor			-0.2571***	(0.0602)	-0.2509***	(0.0682)
Bride's Wealth = Poor, Groom's Wealth = Average			-0.0209	(0.0600)	-0.0138	(0.0663)
Bride's Wealth = Wealthy, Groom's Wealth = Average			0.0101	(0.0609)	0.0342	(0.0671)
Bride's Wealth = Poor, Groom's Wealth = Wealthy			0.0875	(0.0601)	0.0531	(0.0682)
Bride's Wealth = Wealthy, Groom's Wealth = Wealthy			0.2262***	(0.0624)	0.2246***	(0.0713)
Bride More Educated			-0.0483	(0.0428)	-0.0229	(0.0477)
Gov = 1 & Bride's Age = 14					-0.1068	(0.1914)
Gov = 1 & Bride's Age = 15					-0.0608	(0.1738)
Gov = 1 & Bride's Age = 16					-0.2081	(0.1798)
Gov = 1 & Bride's Age = 17					-0.1923	(0.1810)
Gov = 1 & Bride's Age = 18					-0.2279	(0.1763)
Gov = 1 & Bride's Age = 19					-0.2037	(0.1865)
Gov = 1 & Bride's Age = 20					-0.3089*	(0.1841)
Gov = 1 & Bride's Age = 21					-0.3137*	(0.1745)
Gov = 1 & Bride's Age = 22					-0.4170**	(0.1852)
Gov = 1 & Bride's Ed = 8					-0.1095	(0.1192)
Gov = 1 & Bride's Ed = 9					0.0105	(0.1316)
Gov = 1 & Bride's Ed = 10					0.1224	(0.1463)
Gov = 1 & Bride's Ed = 11					0.1619	(0.1471)
Gov = 1 & Bride's Ed = 12					0.0459	(0.1642)
Gov = 1 & Bride's Ed = 13					0.4192***	(0.1622)
Gov = 1 & Bride has Male Friends					0.0486	(0.0773)
Gov = 1 & Dowry					0.0195	(0.0252)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Poor					-0.2537	(0.1584)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Poor					-0.0231	(0.1473)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Average					-0.0616	(0.1585)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Average					-0.0917	(0.1625)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Wealthy					0.1683	(0.1465)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Wealthy					0.0181	(0.1493)
Gov = 1 & Bride More Educated					-0.2075	(0.1351)
Daughter Options	yes		yes		yes	
Match Characteristics	no		yes		yes	
Interactions	no		no		yes	
Number of Choice Experiments	4596		4596		4596	

Notes: Table presents coefficients and standard errors for a probit regression of the following form: $Y_{ir} = 1(\lambda \triangle H_{ir} + \zeta_{ir} > 0)$ where Y_{ir} is equal to 1 if the respondent chose option 1 over option 2 in the groom's side experiment and where $\triangle H_{ir} = H_{ir1} - H_{ir2}$ and H_{irj} gives the characteristics of option $j = \{1, 2\}$ and $\zeta_{ir} \sim IN(0, 1)$. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Determinants of Expected Match

	(1) Education		(2) Government Job		(3) Wealth		(4) Dowry		(5) Dowry	
Age=13	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)
Age=14	0.3276	(0.3188)	0.0586	(0.0724)	0.0379	(0.0728)	-0.3232	(0.2188)	-0.3900*	(0.2013)
Age=15	0.1562	(0.2912)	-0.0435	(0.0622)	0.0302	(0.0663)	-0.1468	(0.1939)	-0.1670	(0.1788)
Age=16	0.3952	(0.2838)	-0.0134	(0.0611)	0.0053	(0.0655)	-0.1423	(0.1906)	-0.1693	(0.1754)
Age=17	0.3439	(0.2838)	0.0469	(0.0594)	-0.0001	(0.0640)	-0.0449	(0.1851)	-0.0931	(0.1700)
Age=18	0.3358	(0.2822)	0.0575	(0.0583)	0.0156	(0.0627)	-0.0190	(0.1830)	-0.1010	(0.1678)
Age=19	0.2987	(0.2827)	0.0034	(0.0586)	0.0074	(0.0629)	0.0717	(0.1831)	-0.0082	(0.1682)
Age=20	0.4383	(0.2779)	0.0636	(0.0579)	0.0324	(0.0624)	-0.0165	(0.1827)	-0.1755	(0.1693)
Age=21	0.4442	(0.2801)	0.0626	(0.0578)	0.0508	(0.0627)	-0.0641	(0.1824)	-0.2701	(0.1727)
Age=22	0.4067	(0.2792)	0.0641	(0.0579)	0.0253	(0.0627)	0.0326	(0.1815)	-0.1837	(0.1726)
Sch=0	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)
Sch=5	1.1546***	(0.2305)	0.0696*	(0.0409)	0.0317	(0.0419)	-0.0099	(0.1262)	-0.1286	(0.1242)
Sch=7	1.7528***	(0.2235)	0.2080***	(0.0408)	0.0799*	(0.0436)	0.2600**	(0.1262)	0.0295	(0.1263)
Sch=8	1.9991***	(0.2174)	0.2235***	(0.0350)	0.0582	(0.0367)	0.1408	(0.1055)	-0.1032	(0.1085)
Sch=9	2.4154***	(0.2164)	0.3220***	(0.0341)	0.1136***	(0.0362)	0.3333***	(0.1060)	-0.0011	(0.1117)
Sch=10	2.7225***	(0.2167)	0.3823***	(0.0343)	0.1112***	(0.0376)	0.4146***	(0.1095)	0.0390	(0.1178)
Sch=11	3.0420***	(0.2160)	0.4320***	(0.0347)	0.1183***	(0.0395)	0.5162***	(0.1144)	0.0918	(0.1239)
Sch=12	3.2783***	(0.2150)	0.4730***	(0.0330)	0.1435***	(0.0374)	0.5835***	(0.1087)	0.1117	(0.1202)
Sch=13	3.3834***	(0.2156)	0.5287***	(0.0342)	0.1797***	(0.0423)	0.7177***	(0.1278)	0.2034	(0.1382)
Wealth of Bride's family=1	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)	0.0000	(.)
Wealth of Bride's family=2	0.0148	(0.0488)	0.0716***	(0.0162)	0.2441***	(0.0185)	1.1221***	(0.0515)	0.9465***	(0.0514)
Wealth of Bride's family=3	0.1573***	(0.0478)	0.1571***	(0.0155)	0.6446***	(0.0194)	2.5414***	(0.0570)	2.0778***	(0.0645)
Daughter Has Male Friends	0.0113	(0.0416)	0.0119	(0.0138)	0.0095	(0.0162)	-0.0381	(0.0504)	-0.0532	(0.0488)
Expect Match with Gov Job									0.2849***	(0.0504)
Expected Age of Match									0.0491***	(0.0147)
Expected Education of Match									0.0723***	(0.0172)
Expected Wealth of Match									0.6260***	(0.0474)
Constant	8.9729***	(0.3287)	0.2369***	(0.0579)	1.8411***	(0.0638)	1.3460***	(0.1784)	-1.5747***	(0.3817)
Observations	4599		4599		4599		4599		4599	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table presents coefficients and standard errors for OLS regressions of characteristics of the 'expected match' given by respondent on characteristics of hypothetical daughter. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Differences in structural parameters by parental education

	Panel A: Preferences									
	Mother has at least primary education					*Father has at least primary education*				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	0.504	(0.133)	0.577	(0.040)	*[0.598]*	0.555	(0.053)	0.590	(0.062)	*[0.666]*
(Age-18) $\times$ 1(Age $\geq$ 18)	-0.299	(0.196)	-0.562	(0.062)	*[0.201]*	-0.446	(0.082)	-0.633	(0.089)	*[0.124]*
Years of Education -7	0.149	(0.068)	0.098	(0.027)	*[0.480]*	0.123	(0.035)	0.090	(0.046)	*[0.573]*
College	-0.663	(0.436)	-0.418	(0.172)	*[0.601]*	-0.603	(0.206)	-0.096	(0.299)	*[0.162]*
Number of experiments	969		5787			3666		2193		
Number of respondents	323		1929			1222		731		
	Panel B: Beliefs									
	Mother has at least primary education					*Father has at least primary education*				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	-0.004	(0.013)	-0.019	(0.008)	*[0.330]*	0.000	(0.011)	0.000	(0.013)	*[0.979]*
Education	1.474	(0.285)	1.156	(0.175)	*[0.341]*	1.242	(0.137)	0.809	(0.527)	*[0.427]*
College	3.141	(0.738)	3.963	(0.453)	*[0.343]*	3.551	(0.421)	4.963	(0.677)	*[0.077]*
Age $\times$ Education	-0.123	(0.026)	-0.098	(0.016)	*[0.408]*	-0.099	(0.011)	-0.072	(0.051)	*[0.594]*
Constant	-3.377	(0.135)	-3.227	(0.075)	*[0.332]*	-3.412	(0.108)	-3.401	(0.119)	*[0.944]*
	878		4487			1212		715		
	346		1844			2996		1730		

Notes: Table presents preference and belief structural parameters related to a daughters age at marriage and education, estimated separately by parental education. In doing so, we hold all other preference parameters constant at the values in Column (3) of Table II and all other belief parameters constant at the values in Table III. p-values are testing the null hypothesis that the coefficients are the same across groups. p-values and standard errors (in parentheses) estimated using a bootstrap, resampling respondents with replacement.

Table A.8: Differences in structural parameters by parental socio-economic status

	Panel A: Preferences									
	Scheduled Caste or Tribe					***Dirt floor***				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	0.459	(0.051)	0.637	(0.059)	*[0.023]*	0.473	(0.047)	0.687	(0.069)	*[0.010]*
(Age-18) $\times$ 1(Age $\geq$ 18)	-0.407	(0.082)	-0.587	(0.085)	*[0.129]*	-0.430	(0.077)	-0.635	(0.094)	*[0.092]*
Years of Education -7	0.108	(0.041)	0.100	(0.034)	*[0.875]*	0.108	(0.035)	0.105	(0.037)	*[0.938]*
College	-0.254	(0.282)	-0.570	(0.200)	*[0.361]*	-0.322	(0.222)	-0.650	(0.223)	*[0.298]*
Number of experiments	2511		4425			3636		3333		
Number of respondents	837		1475			1212		1111		
	Panel B: Beliefs									
	Scheduled Caste or Tribe					***Dirt floor***				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	-0.001	(0.016)	-0.007	(0.013)	*[0.777]*	-0.013	(0.007)	0.000	(0.010)	*[0.334]*
Education	0.951	(0.476)	1.445	(0.106)	*[0.312]*	0.768	(0.463)	1.332	(0.091)	*[0.232]*
College	4.598	(0.600)	3.494	(0.251)	*[0.090]*	4.647	(0.521)	3.225	(0.337)	*[0.022]*
Age $\times$ Education	-0.083	(0.046)	-0.127	(0.012)	*[0.349]*	-0.060	(0.047)	-0.106	(0.008)	*[0.326]*
Constant	-3.406	(0.131)	-3.347	(0.128)	*[0.747]*	-3.270	(0.077)	-3.381	(0.103)	*[0.392]*
	1869		3624			2671		2851		
	769		1477			1110		1147		

Notes: Table presents preference and belief structural parameters related to a daughters age at marriage and education, estimated separately by parental socio-economic status. In doing so, we hold all other preference parameters constant at the values in Column (3) of Table II and all other belief parameters constant at the values in Table III. p-values are testing the null hypothesis that the coefficients are the same across groups. p-values and standard errors (in parentheses) estimated using a bootstrap, resampling respondents with replacement.

Table A.9: Differences in structural parameters by whether respondent has a married child (son or daughter) and by the age of the daughter

	Panel A: Preferences									
	Respondent has at least one married child					***Respondent's daughter is older***				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	0.565	(0.060)	0.572	(0.049)	*[0.932]*	0.612	(0.060)	0.530	(0.051)	*[0.292]*
(Age-18) $\times$ 1(Age $\geq$ 18)	-0.483	(0.092)	-0.550	(0.076)	*[0.576]*	-0.535	(0.083)	-0.503	(0.084)	*[0.784]*
Years of Education -7	0.053	(0.041)	0.137	(0.032)	*[0.109]*	0.132	(0.034)	0.062	(0.038)	*[0.171]*
College	-0.683	(0.243)	-0.338	(0.204)	*[0.278]*	-0.703	(0.212)	-0.143	(0.244)	*[0.083]*
Number of experiments	2553		4374			3729		3243		
Number of respondents	851		1458			1243		1081		
	Panel B: Beliefs									
	Respondent has at least one married child					***Respondent's daughter is older***				
	Yes		**No**		***p***	**Yes**		**No**		***p***
Age	-0.001	(0.012)	-0.001	(0.012)	*[0.964]*	-0.001	(0.012)	-0.003	(0.009)	*[0.903]*
Education	1.343	(0.135)	1.165	(0.358)	*[0.642]*	1.018	(0.335)	1.249	(0.302)	*[0.608]*
College	3.392	(0.426)	3.994	(0.533)	*[0.378]*	4.148	(0.557)	3.289	(0.654)	*[0.317]*
Age $\times$ Education	-0.110	(0.012)	-0.099	(0.034)	*[0.757]*	-0.082	(0.033)	-0.096	(0.028)	*[0.746]*
Constant	-3.405	(0.121)	-3.395	(0.117)	*[0.956]*	-3.400	(0.120)	-3.385	(0.091)	*[0.920]*
	1946		3548			2934		2590		
	802		1444			1208		1050		

Notes: Table presents preference and belief structural parameters related to a daughters age at marriage and education, estimated separately by: (1) whether or not the respondent has a married child (daughter or son) and by (2) whether or not they have a daughter in the "older" cohort (aged 15-17 at baseline). In doing so, we hold all other preference parameters constant at the values in Column (3) of Table II and all other belief parameters constant at the values in Table III. p-values are testing the null hypothesis that the coefficients are the same across groups. p-values and standard errors (in parentheses) estimated using a bootstrap, resampling respondents with replacement.

Table A.10: Differences in structural parameters by RCT treatment status

	Panel A: Preferences *By Treatment Status*							
	Control		**Treatment 1**		***p***	**Treatment 2**		***p***
Age	0.625	(0.068)	0.554	(0.075)	*[0.483]*	0.511	(0.056)	*[0.202]*
(Age-18) $\times$ 1(Age $\geq$ 18)	-0.479	(0.102)	-0.484	(0.112)	*[0.975]*	-0.569	(0.089)	*[0.505]*
Years of Education -7	0.090	(0.040)	0.120	(0.042)	*[0.604]*	0.119	(0.043)	*[0.623]*
College	-0.630	(0.291)	-0.383	(0.263)	*[0.529]*	-0.426	(0.261)	*[0.601]*
Number of experiments	2418		2205			2349		
Number of respondents	806		735			783		

	Panel B: Beliefs *By Treatment Status*							
	Control		**Treatment 1**		***p***	**Treatment 2**		***p***
Age	-0.014	(0.012)	-0.007	(0.010)	*[0.684]*	-0.015	(0.011)	*[0.925]*
Education	1.288	(0.105)	0.902	(0.419)	*[0.371]*	1.009	(0.319)	*[0.406]*
College	3.378	(0.353)	4.283	(0.631)	*[0.211]*	3.789	(0.560)	*[0.534]*
Age $\times$ Education	-0.103	(0.010)	-0.070	(0.040)	*[0.422]*	-0.073	(0.031)	*[0.354]*
Constant	-3.277	(0.125)	-3.317	(0.100)	*[0.803]*	-3.257	(0.110)	*[0.902]*
	2008		1757			1759		
	813		719			726		

Notes: Table presents preference and belief structural parameters related to a daughters age at marriage and education, estimated separately by treatment status in the RCT. In doing so, we hold all other preference parameters constant at the values in Column (3) of Table II and all other belief parameters constant at the values in Table III. p-values are testing the null hypothesis that the coefficients are the same across groups. p-values and standard errors (in parentheses) estimated using a bootstrap, resampling respondents with replacement. RCT treatment arms include: control, Treatment 1 (working with girls only) and Treatment 2 (working with girls and their communities). Details of intervention given in Andrew et al. (2022).

Table A.11: Testing the Impact of Vignette Salience on Response Patterns

	Panel A: Ex-Post Experiment			
	(1)	(2)	(3)	(4)
	Probit coefs.	Marginal Effects	Probit coefs.	Marginal Effects
Salient option	0.0347	0.0124	0.0425	0.0148
	(0.0373)	(0.0133)	(0.0379)	(0.0132)
Age x Education Interactions	yes		yes	
Groom/Match characteristics	no		yes	
Number of Choice Experiments	6972		6972	
Number of Respondents	2324		2324	
	Panel B: Ex-Ante Experiment			
	(1)	(2)	(3)	(4)
	Probit coefs.	Marginal Effects	Probit coefs.	Marginal Effects
Salient vignette	-0.0070	-0.0022	-0.0139	-0.0044
	(0.0432)	(0.0138)	(0.0435)	(0.0137)
Age x Education Interactions	yes		yes	
Groom/Match characteristics	no		yes	
Number of Choice Experiments	5524		5524	
Number of Respondents	2258		2258	

Notes: Table presents coefficients and Average Marginal Effects (and standard errors thereof) for the β coefficients from the following probit regressions. Panel A: $Y_{ir} = 1\left(\beta Salient_{ir} + \lambda \triangle H_{ir} + \zeta_{ir} > 0\right)$ where Y_{ir} is equal to 1 if the respondent chose option 1 over option 2 in round r of the ex-post experiment, where $Salient_{ir}$ indicates that option 1 in round r was the salient option and where $\triangle H_{ir} = H_{ir1} - H_{ir2}$ and H_{irj} gives the characteristics of option $j = \{1, 2\}$ and $\zeta_{ir} \sim IN(0, 1)$. Panel B: $Y_{ir} = 1\left(\beta Salient_{ir} + \lambda H_{ir} + \zeta_{ir} > 0\right)$ where Y_{ir} is equal to 1 if the respondent chose to marry the daughter in round r of the ex-ante experiment, where $Salient_{ir}$ indicates that round r was the salient vignette and where H_{ir} gives the characteristics of round r and $\zeta_{ir} \sim IN(0, 1)$. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: Ex Post Experiment: Example Script

Background
Please imagine a mother and father named Ramesh and Rita, they live in a village similar to your village. Ramesh and have a daughter named Sita.

Compared to other households in the village, Ramesh and Rita's household is of average wealth. Their house has two rooms with a dung floor. They own one bigha of land and two cows. They own an electric fan and a bicycle but not a TV.

Sita's age is 12. Sita's currently studying in 7th standard. Sita is getting average grades at school and really enjoys going. At Sita's school there is a scheme that covers all the costs of Sita's education, including stationary, uniforms and transport.

Sita's mother, Rita, really struggles to take care of all the work that needs doing at home. So Sita's help with cooking, cleaning, and taking care of elderly relatives is very useful.

Ramesh and Rita are considering when and to whom they will get Sita married and until when they will keep her in education. Imagine there are two possible options, for when Sita will leave education and when and to whom she will get married.

Option 1
Sita completes standard 9 at school. Sita marries at age 16.

She marries Rahul. Rahul is 24 years old. Rahul attended College and has a government job.

Compared to other household in the village, Rahul's household is very wealthy. Their house has three rooms with a cement floor. They own four bighas of land and two cows. As well as an electric fan and a TV they also own a refrigerator and a motorcycle. Rahul's parents expect at least 6 lakhs in marriage gifts.

Option 2
Sita completes standard 12 at school. Sita marries at age 19.

She marries Bharat. Bharat is 22 years old. Bharat attended school until 10th standard.

Compared to other household in the village, Bharat's household is very poor. The whole household live in one room with a dung floor and they don't own any land. They have one cow. They own very other few assets, for example, they don't own an electric fan, a TV or a bicycle. Bharat's parents expect at least 2.5 lakhs in marriage gifts.

Choice
Which option do you think Ramesh and Rita will choose for their daughter?

1. Keep Sita in education until she has finished standard 9 and then marry her to Rahul when she is age 16

2. Keep Sita in education until she has finished standard 12 and then marry her to Bharat when she is age 19

Figure A.2: Ex Ante Experiment: Example Script

Background
Please imagine a mother and father named Raj Kumar and Aarti, they live in a village similar to your village. Raj Kumar and Aarti have a daughter named Jyoti.

Compared to other households in the village, Raj Kumar and Aartis' household is of average wealth. Their house has two rooms with a dung floor. They own one bigha of land and two cows. They own an electric fan and a bicycle but not a TV.

Jyoti's age is 15. She is currently studying in 10th standard.

Jyoti is getting average grades at school but does not enjoy school. Jyoti's parents have to pay for the full cost of Jyoti's education, including stationary, uniforms and transport.

Jyoti's family have no particular need for Jyoti to spend lots of time helping at home. When Jyoti is at home she spends lots of her time sitting and relaxing.

Raj Kumar and Aart are worried about Jyoti as she is friends with some boys and sometimes stays out of the house until late.

Offer in Hand
Raji Kuma and Aarti are considering whether they will get Jyoti married in the next year, whether they will keep her in school for another year or whether Jyoti will leave school to help her parents at home.

Raj Kumar and Aarti know of a potential suitor for Jyoti, Amit.

Compared to other households in the village, Amit's household is very poor. The whole household live in one room with a dung floor and they don't own any land. They have one cow. They own very other few assets, for example, they don't own an electric fan, a TV or a bicycle.

Amit's parents expect at least 4 lakhs in marriage gifts.

Amit is 24 years old. Amit attended school until 11th standard.

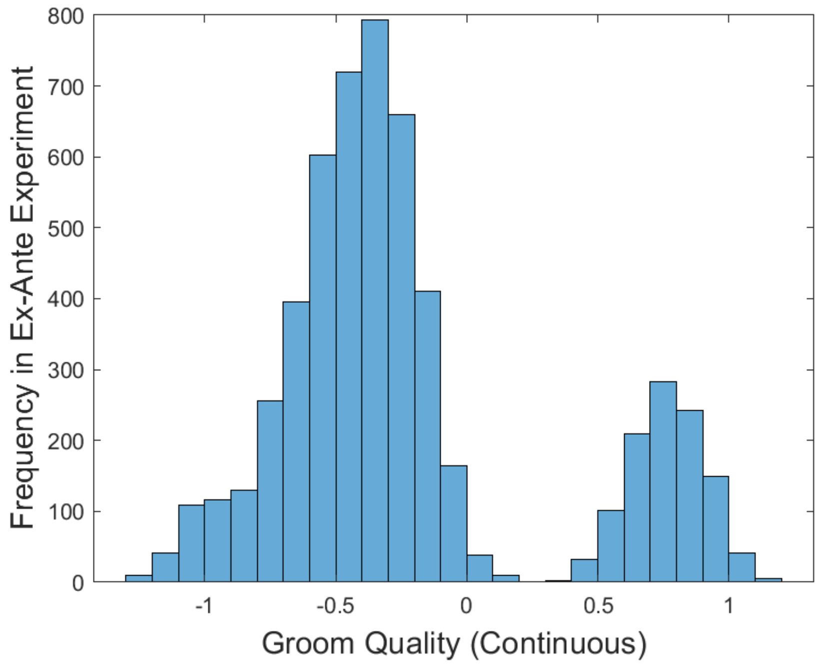
Choice
Which option do you think Raj Kumar and Aarti will choose for their daughter?

1. Marry Jyoti to Amit this year

2. Keep Jyoti in school this year

3. Take Jyoti out of school so she can help in the home this year

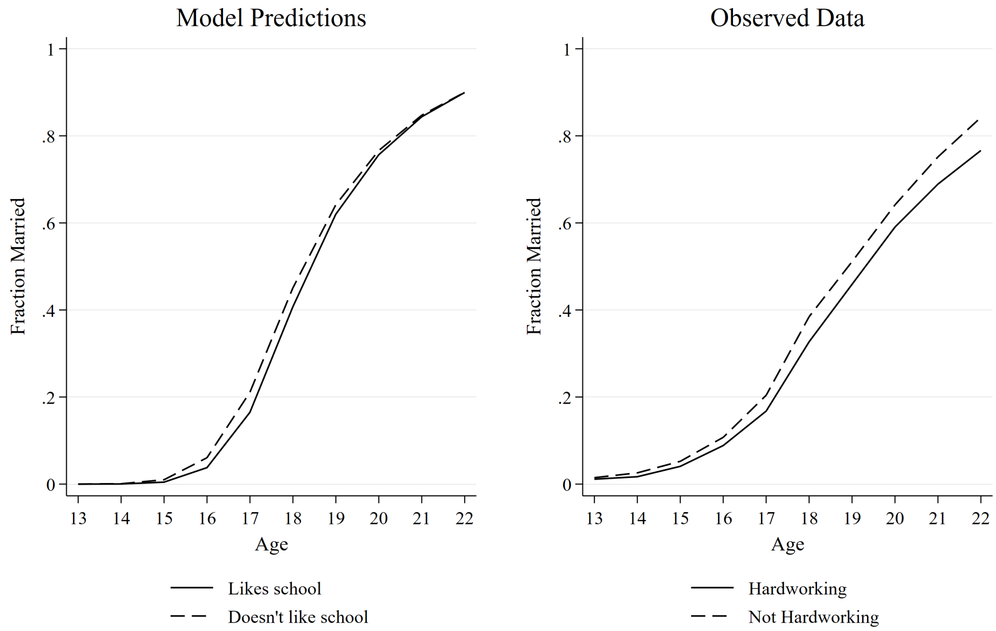
Figure A.3: Histogram of groom quality



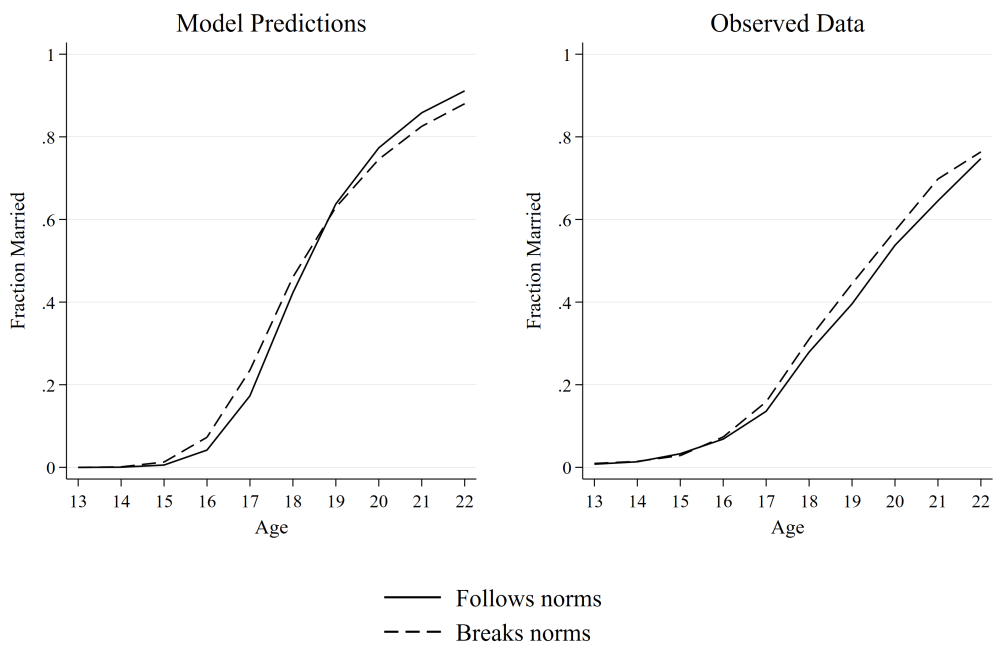
Notes: Figure plots histogram of estimated groom quality index over the potential grooms that we presented to respondents in the ex-ante experiments. For a groom with characteristics X^g, their quality index is constructed as $X^g\hat{\beta}^g$ where β^g is the vector of estimated preference parameters on groom characteristics at Table II column 3.

Figure A.4: Implied vs. Observed heterogeneity in marriage by shifters of preferences and beliefs

(a) Heterogeneity by whether or not girl likes school (in the choice experiments) and whether or not mother reported girl is hardworking (observational data)

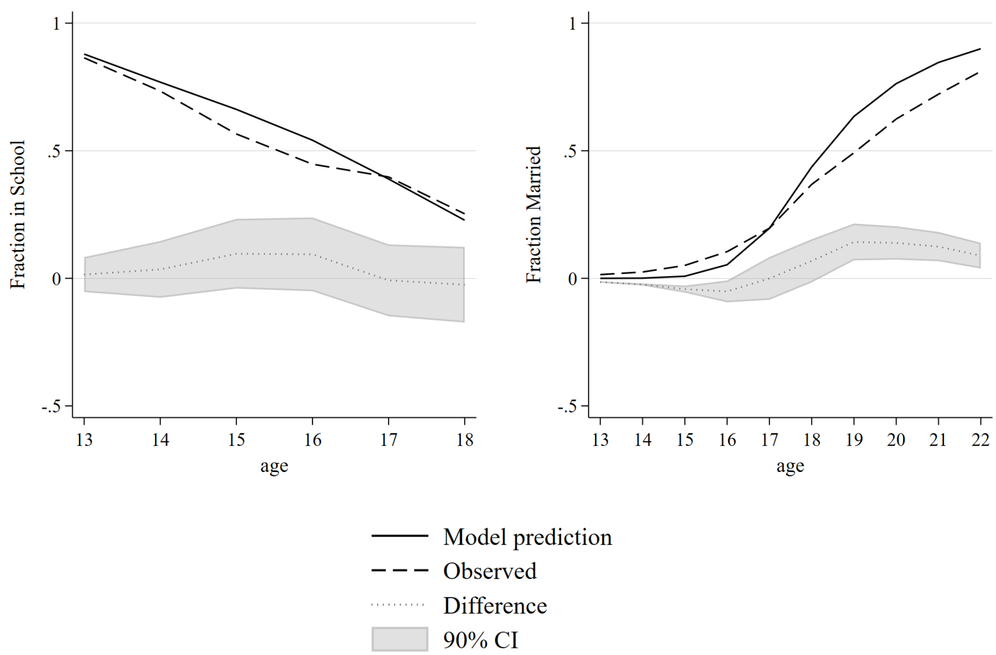


(b) Heterogeneity by whether or not parents are worried that daughters behavior conflicts with gender norms (gets home late and has unsuitable friends)



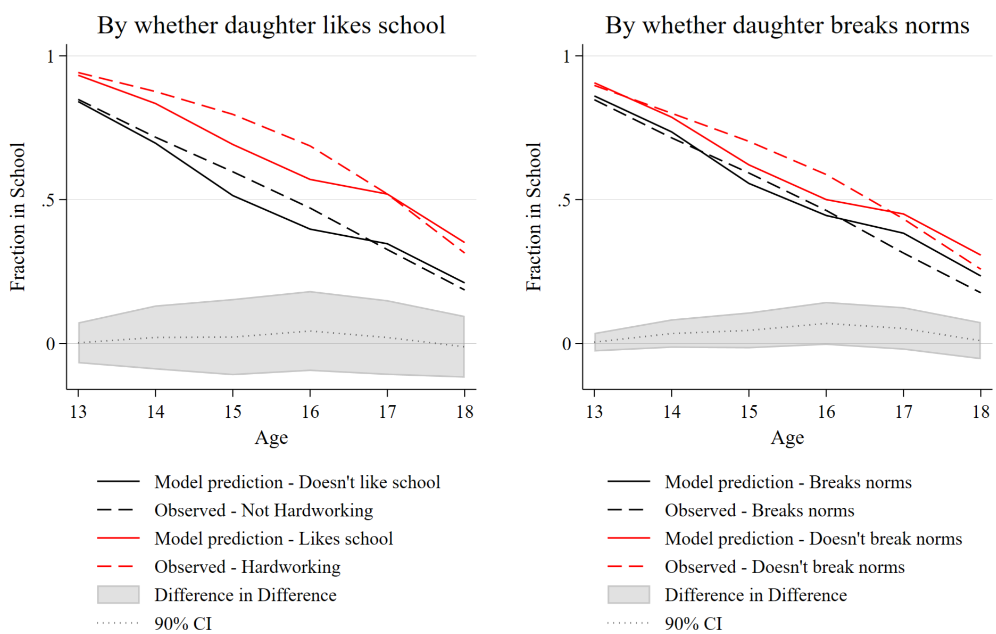
Notes: Figures plot the predicted (left) and observed (right) proportion of girls married at each age, split by: (a) whether or not a daughter likes school (predicted) or is hardworking (observed); and (b) whether or not a daughter follows gender norms.

Figure A.5: Formal comparison between model predictions and observed patterns



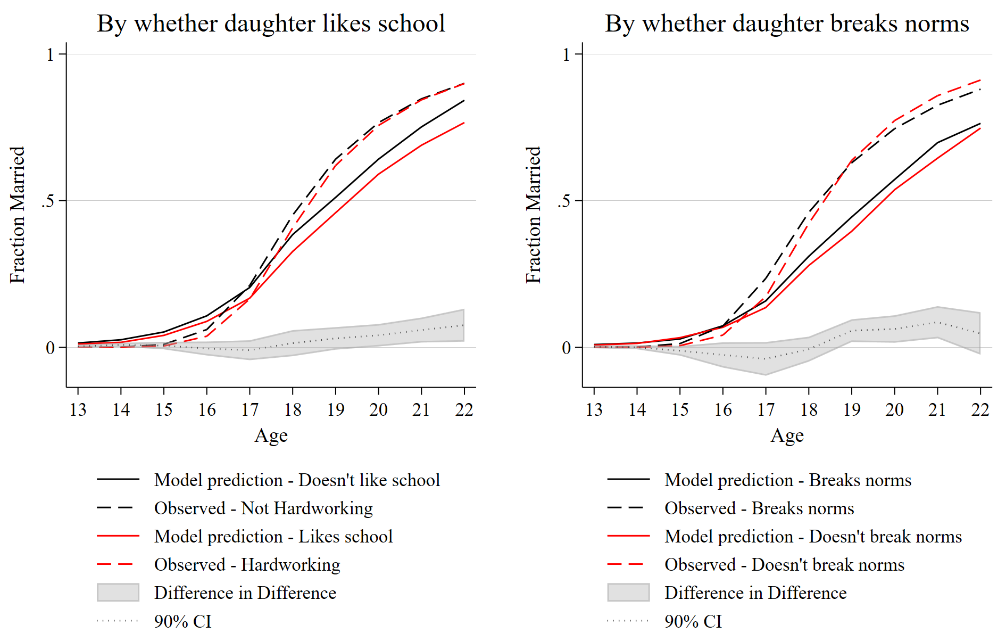
Notes: Figures plot the proportion in school (left) and married (right) by age as predicted by our model estimates (solid lines) and as observed in the observational data (dashed lines). We also plot the difference between these values by age (model predictions - observed) and the 90% confidence interval around this difference. We construct this confidence interval accounting for both uncertainty in the model estimates (by creating predicted trajectories for every bootstrapped sample) and uncertainty in the observational estimates.

Figure A.6: Formal comparison between model predictions and observed patterns for schooling by shifters of preferences and beliefs



Notes: Figures plot the proportion in school by age as predicted by our model estimates (solid lines) and as observed in the observational data (dashed lines). In these figures we split by whether or not the daughter in question likes school/is hardworking (left) and whether or not the daughter in question breaks with gendered norms of behavior (right). We also plot the difference in differences between these values by age ($(M^1 - D^1) - (M^0 - D^0)$) where M^k gives the model prediction for $k = \{0, 1\}$ and D^k gives the value observed in the data. k denotes the heterogeneity split. On the left hand side, we split by whether a daughter likes school ($k = 1$, in red) or doesn't like school ($k = 0$, in black). We also plot the 90% confidence interval around this difference in difference. We construct this confidence interval accounting for both uncertainty in the model estimates (by creating predicted trajectories for every bootstrapped sample) and uncertainty in the observational estimates.

Figure A.7: Formal comparison between model predictions and observed patterns for marriage by shifters of preferences and beliefs



Notes: Figures plot the proportion married by age as predicted by our model estimates (solid lines) and as observed in the observational data (dashed lines). In these figures we split by whether or not the daughter in question likes school/is hardworking (left) and whether or not the daughter in question breaks with gendered norms of behavior (right). We also plot the difference in differences between these values by age ($(M^1 - D^1) - (M^0 - D^0)$) where M^k gives the model prediction for $k = \{0, 1\}$ and D^k gives the value observed in the data. k denotes the heterogeneity split. On the left hand side, we split by whether a daughter likes school ($k = 1$, in red) or doesn't like school ($k = 0$, in black). We also plot the 90% confidence interval around this difference in difference. We construct this confidence interval accounting for both uncertainty in the model estimates (by creating predicted trajectories for every bootstrapped sample) and uncertainty in the observational estimates.

Figure A.8: Validation Experiments: Expected Match Visual Aid

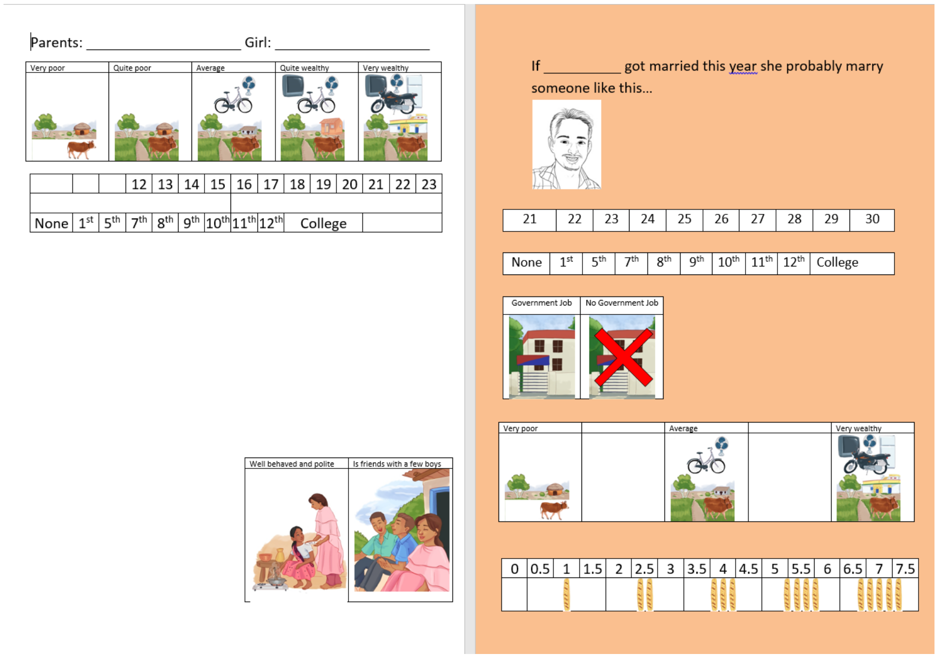
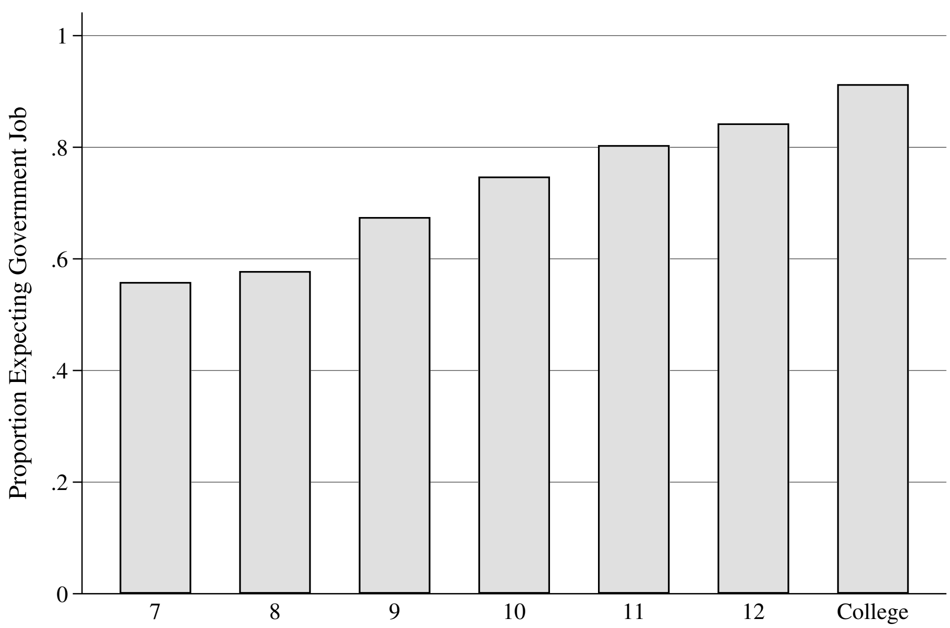


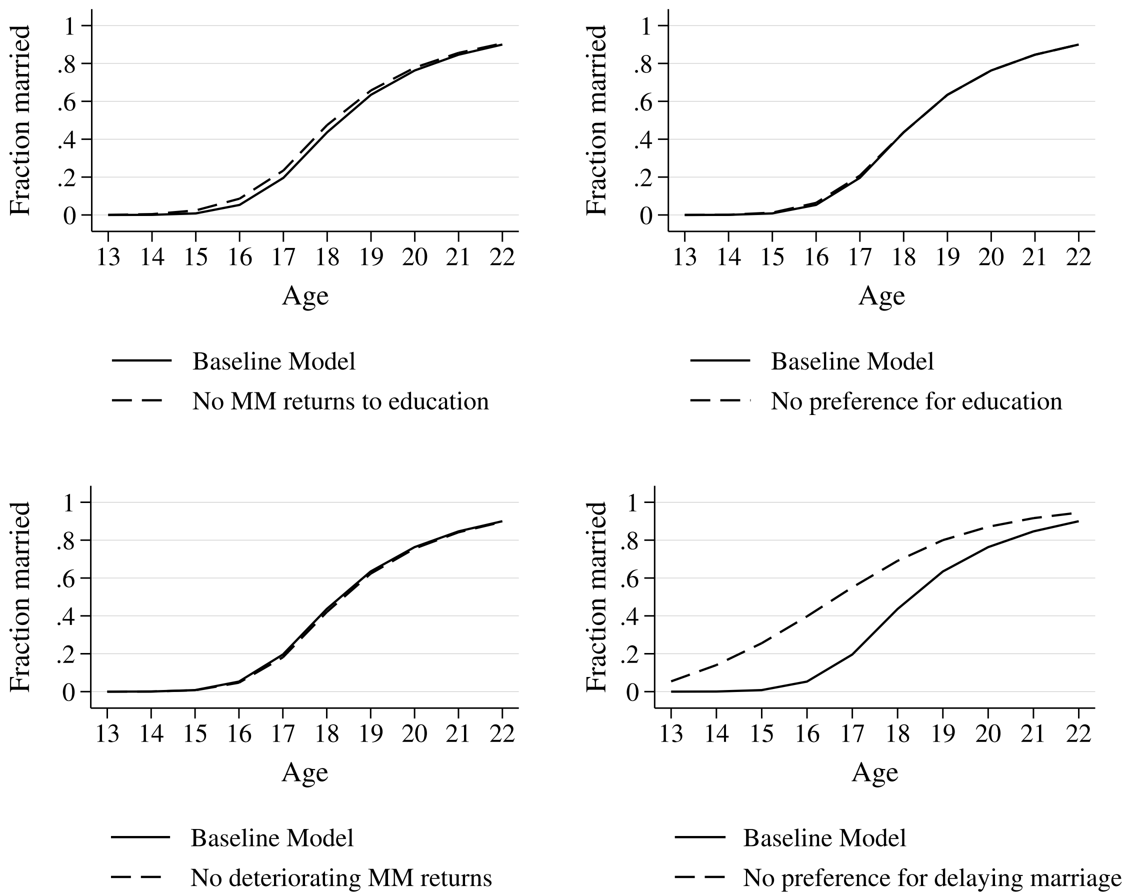
Figure A.9: Expected Match Has a Government Job by Girl's Education



Notes: Bar chart plots the fraction of times respondents answered that they expected a hypothetical daughter would marry a groom with a government job if she got married this year, split by the education of the hypothetical daughter.

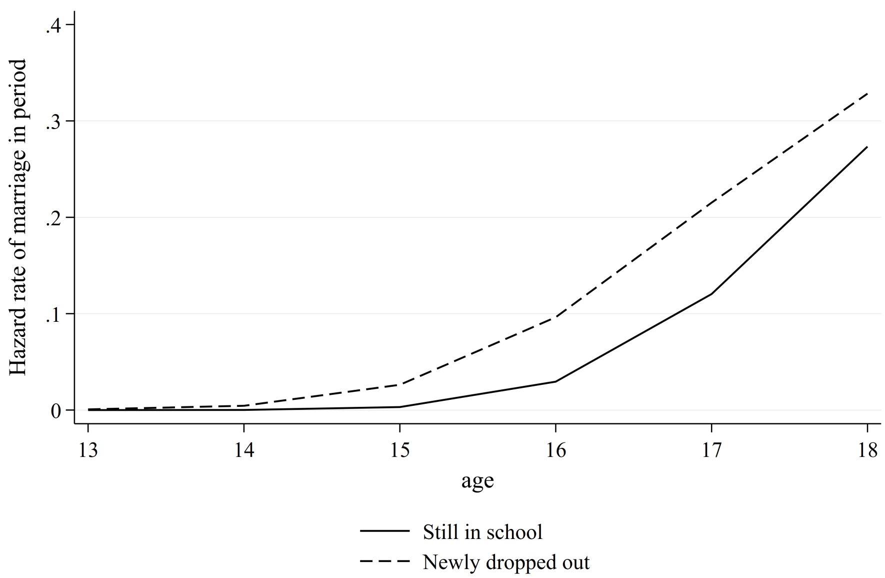
Figure A.10: Validation Experiments: Groom Experiment Visual Aid

Figure A.11: Counterfactual marriage profiles



Notes: Graphs show baseline and counterfactual patterns for the proportion of girls married by age for four different counterfactuals: (i) Parents perceive no marriage market return to education; (ii) Parents have no preference for a daughter's education; (iii) Parents perceive no deteriorating marriage market returns to age; (iv) Parents have no preference for delaying marriage.

Figure A.12: Impact of shocks on hazard rate of marriage by age



Notes: Graphs show baseline and counterfactual hazard rate for marriage by age for girls who are still in school vs. newly dropped out.

B Identification

B.1 A Note on Timing

We define all utility profiles and value functions from the perspective of period $t = 0$. We do this to avoid taking a stance on the allocation of payoffs across periods.[72] This means that conditional value function $v(\bar{\omega}_t, q_t, d_t = d)$ is the discounted value of being in period 0 but knowing with certainty that period t's state variables will be $\{\bar{\omega}_t, q_t\}$ and period t's choice will be $d_t = d$. All information needed to form payoffs that occur between period 0 and $t-1$ can be reconstructed from the period t deterministic state variables. This means that there is a straightforward relationship between value functions defined from the point of view of period 0 and those defined from the point of view of period t ($v_t(\bar{\omega}_t, q_t, d_t = d)$). In particular:

$$v(\bar{\omega}_t, q_t, d_t = d) = \sum_{\tau=0}^{t-1} \beta^\tau u(\bar{\omega}_\tau|_{\bar{\omega}_t}, d_\tau|_{\omega t}) + \beta^t v_t(\bar{\omega}_t, q_t, d_t = d)$$

where $\bar{\omega}_\tau|_{\bar{\omega}_t}$ and $d_\tau|_{\boldsymbol{\omega}_t}$ are the state variables and the actions at τ implied by $\bar{\omega}_t$ and $u(\bar{\omega}_\tau|_{\bar{\omega}_t}, d_\tau|_{\omega t})$ are the flow payoffs accrued in period τ.

Working with value functions defined from the perceptive period t implies the same choice probabilities in the ex-ante experiment as those defined from the perceptive period 0 and used in equation (11) :

$$\begin{aligned} p^A(d_t = d|\bar{\omega}_t, q_t) &= \mathbb{P}\left[v_t(\bar{\omega}_t, q_t, d_t = d) + \varepsilon_{i,t}(d) > v_t(\bar{\omega}_t, q_t, d_t = d') + \varepsilon_{i,t}(d') \quad \forall d' \neq d \quad \right] \\ &= \mathbb{P}\left[v(\bar{\omega}_t, q_t, d_t = d) + \beta^t \varepsilon_{i,t}(d) > v(\bar{\omega}_t, q_t, d_t = d') + \beta^t \varepsilon_{i,t}(d') \quad \forall d' \neq d \quad \right] \end{aligned}$$

This structure does imply a particular form of heteroskedasticity in the ex-ante experiment. Namely, the variance of errors should be decreasing in age.

B.2 Proof of Theorem 1.

Let $|\bar{\Psi}| = M$ and $\mathbf{P}^P$ give the $M(M-1)/2$-vector of ex-post choice probabilities (that the first option is chosen) at the possible combinations of paths presented in the ex-post experiment. Inverting the expression for ex-post choice probabilities (Equation 7) and leveraging our assumption on the error term gives:

$$\sqrt{2}\Phi^{-1}(\mathbf{P}^P) = BU \tag{20}$$

[72] Taking education, as an example, this choice means we avoid having to take a stance on whether parents receive flow payoffs from a daughter in school or whether they receive a payoff dependent on the total amount of education she accrues.

where

$$\Phi^{-1}(\mathbf{P}^P) = \begin{bmatrix} \Phi^{-1}(p^P(\bar{\Psi}^0, \bar{\Psi}^1)) \\ \Phi^{-1}(p^P(\bar{\Psi}^0, \bar{\Psi}^2)) \\ \vdots \\ \Phi^{-1}(p^P(\bar{\Psi}^{M-1}, \bar{\Psi}^M)) \end{bmatrix} \quad , \quad U = \begin{bmatrix} 1/\sigma_\nu \\ \bar{U}(\bar{\Psi}^2)/\sigma_\nu \\ \vdots \\ \bar{U}(\bar{\Psi}^M)/\sigma_\nu \end{bmatrix}$$

and

$$B = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & -1 \end{bmatrix}$$

That is, $\Phi^{-1}(\mathbf{P}^P)$ is an $M(M-1)/2$-vector of the quantiles of the normal distribution at the observed choice probabilties, U is the $M-1$ vector of utility values (relative to the path $\bar{\Psi}^0$ and scaled by the difference $\bar{U}(\bar{\Psi}^1) - \bar{U}(\bar{\Psi}^0)$ which we normalize to 1), and B is the design matrix that gives every unique combination of utility differences. Pre-multiplying both sides of Equation 20 by B' and then inverting we get:

$$U_0 = \sqrt{2}(B'B)^{-1}B'\Phi^{-1}(\mathbf{P}^P) \tag{21}$$

Thus, utility relative to path $\bar{\Psi}^0$ can be identified from the observed ex-post choice probabilities.

B.3 Proof of Theorem 2.

Step 1. Recover conditional value functions. In Step 1, we invert equation (11) in the main text to recover $v(\bar{\omega}_t, q_t, d_t)$ at all $\bar{\omega}_t, q_t, d_t$.

$$p^A(d_t = d|\bar{\omega}_t, q_t) = \mathbb{P}\left[v(\bar{\omega}_t, q_t, d_t = d) + \beta^t \varepsilon_{i,t}(d) > v(\bar{\omega}_t, q_t, d_t = d') + \beta^t \varepsilon_{i,t}(d') \quad \forall d' \neq d \quad \right]$$

$p^A(d_t = d|\bar{\omega}_t, q_t)$ is observed directly from our ex-ante experiment at all $\{d_t, \bar{\omega}_t, q_t\}$. We need to identify $v(\bar{\omega}_t, q_t, d_t)$ for all $d_t, \bar{\omega}_t, q_t$ and σ^2_ε. This would seem to be impossible as the order condition appears to be violated. However, the terminal action provides restrictions on $v(\bar{\omega}_t, q_t, d_t)$ which reduce the number of independent objects that need identifying. Namely, when the terminal action is chosen ($d_t = 0$), the conditional value function is the utility of the completed path: $v(q_t, \bar{\omega}_t, d_t = 0) \equiv \bar{U}(t, q_t, \bar{\omega}_t)$. Note that this linkage is what provides the terminal location and scaling normalizations in this problem. It also means that we only need to identify $v(q_t, \bar{\omega}_t, d_t)$ for $d_t > 0$. Further, under our assumption of the irrelevance of rejected offers, $v(L, \bar{\omega}_t, d_t) = v(H, \bar{\omega}_t, d_t)$ for all $\bar{\omega}_t, d_t > 0$.

There are $D + 1$ possible actions. For every $\bar{\omega}_t$, we have $2 \times D$ independent expressions (i.e. D for each of the two quality levels) and we have $D + 1$ unknowns (σ_ε^2 and $v(., \bar{\omega}_t, d_t)$ for each $d_t = 1, ..D$). The order and rank conditions hold whenever $D \geq 1$ (i.e. in any meaningful binary or discrete choice) and when $U(t, L, \bar{\omega}_t) \neq U(t, H, \bar{\omega}_t)$ for all $\{t, \bar{\omega}_t\}$.

For $D > 1$, no closed form solutions exist but, given our normality assumption, identification proceeds analogously. Convenient closed form expressions exist in the binary, $D = 1$, case as shown below where we use the fact that $v(\bar{\omega}_t, H, d_t = 1) = v(\bar{\omega}_t, L, d_t = 1)$.

$$p^A(d_t = 0|\bar{\omega}_t, L) = \Phi\left(\frac{\bar{U}(t, L, \bar{\omega}_t) - v(\bar{\omega}_t, L, d_t = 1)}{\beta^t\sqrt{2}\sigma_\varepsilon}\right) \tag{22}$$

$$p^A(d_t = 0|\bar{\omega}_t, H) = \Phi\left(\frac{\bar{U}(t, H, \bar{\omega}_t) - v(\bar{\omega}_t, L, d_t = 1)}{\beta^t\sqrt{2}\sigma_\varepsilon}\right) \tag{23}$$

Inverting and rearranging, we can recover σ_ε and $v(\bar{\omega}_t, L, d_t = 1)$.

Step 2. Exploiting the recursive relationship between value functions. Combining equations (9) and (10) gives the following recursive relationship between conditional value functions whenever the terminal action is not chosen (i.e. $d_t > 0$):[73]

$$
\begin{aligned}
&v(\bar{\omega}_t, q_t, d_t = 1) \\
&\quad = \pi(\bar{\omega}_{t+1}) \sum_{d \in O(\bar{\omega}_{t+1})} p^A(d_{t+1} = d|\bar{\omega}_{t+1}, H)\left[v(\bar{\omega}_{t+1}, H, d_{t+1} = d) + \beta E(\varepsilon_{t+1}(d)|d_{t+1} = d, \bar{\omega}_{t+1}, H)\right] \\
&\quad + (1 - \pi(\bar{\omega}_{t+1})) \sum_{d \in O(\bar{\omega}_{t+1})} p^A(d_{t+1} = d|\bar{\omega}_{t+1}, L)\left[v(\bar{\omega}_{t+1}, L, d_{t+1} = d) + \beta E(\varepsilon_{t+1}(d)|d_{t+1} = d, \bar{\omega}_{t+1}, L)\right] \\
&\quad \text{s.t. } \bar{\omega}_{t+1} = \bar{\omega}'(\bar{\omega}_t, d_t)
\end{aligned} \tag{25}
$$

Although there are many terms here, the key point to note is that at this stage $\pi(\bar{\omega}_{t+1})$ is the only remaining unknown since: (i) $p^A(d_{t+1} = d|\bar{\omega}_{t+1}, q)$ is directly observed from the choice experiments for all $\bar{\omega}_{t+1}, q$; (ii) $v(\bar{\omega}_{t+1}, q, d_{t+1} = d)$ was identified in step 1 for all $\bar{\omega}_{t+1}, q, d$; (iii) $E(\varepsilon_{t+1}(d)|d_{t+1} = d, \bar{\omega}_{t+1}, q)$, i.e. the expected value of idiosyncratic preference shocks conditional on optimal actions, can be straightforwardly calculated.[74] This allows us to insert expressions for (i), (ii) and (iii) into equation (25) for each $\bar{\omega}_t$ and rearrange constructively to identify $\pi(\bar{\omega}_t)$.

[73] In the binary choice ($D = 1$) version of the problem, this relation is as follows:

$$
\begin{aligned}
&v(\bar{\omega}_t, q_t, d_t) \\
&= \underbrace{(1 - \pi(\bar{\omega}_{t+1}))p^A(d_{t+1} = 0|\bar{\omega}_{t+1}, L)\left[v(\bar{\omega}_{t+1}, L, d_{t+1} = 0) + \beta \mathbb{E}\left(\varepsilon_t(0)|\bar{\omega}_{t+1}, L, d^*_{t+1} = 0\right)\right]}_{\text{low offer, accept}} \\
&+ \underbrace{\pi(\bar{\omega}_{t+1})p^A(d_{t+1} = 0|\bar{\omega}_{t+1}, H)\left[v(\bar{\omega}_{t+1}, H, d_{t+1} = 0) + \beta \mathbb{E}\left(\varepsilon_t(0)|\bar{\omega}_{t+1}, H, d^*_{t+1} = 0\right)\right]}_{\text{high offer, accept}} \\
&+ \underbrace{(1 - \pi(\bar{\omega}_{t+1}))p^A(d_{t+1} = 1|\bar{\omega}_{t+1}, L)\left[v(\bar{\omega}_{t+1}, L, d_{t+1} = 1) + \beta \mathbb{E}\left(\varepsilon_t(1)|\bar{\omega}_{t+1}, L, d^*_{t+1} = 1\right)\right]}_{\text{low offer, reject}} \\
&+ \underbrace{\pi(\bar{\omega}_{t+1})p^A(d_{t+1} = 1|\bar{\omega}_{t+1}, H)\left[v(\bar{\omega}_{t+1}, H, d_{t+1} = 1) + \beta \mathbb{E}\left(\varepsilon_t(1)|\bar{\omega}_{t+1}, H, d^*_{t+1} = 1\right)\right]}_{\text{high offer, reject}}
\end{aligned} \tag{24}
$$

[74] When the choice set is binary ($D = 1$), these have closed form expressions. For example:

$$
\mathbb{E}\left(\varepsilon_t(0)|\bar{\omega}_t, q_t, d^\star_t = 0\right) = \mathbb{E}\left(\varepsilon_t(0)|\bar{U}(t, q_t, \bar{\omega}_t) + \beta^t \varepsilon_t(0) \geq v(\bar{\omega}_t, q_t, d_t = 1) + \beta^t \varepsilon_t(1)\right) \tag{26}
$$

$$
= \mathbb{E}\left(\varepsilon_t(0)|\varepsilon_t(0) - \varepsilon_t(1) \geq \frac{v(\bar{\omega}_t, q_t, d_t = 1) - \bar{U}(t, q_t, \bar{\omega}_t)}{\beta^t}\right) \tag{27}
$$

$$
= \frac{1}{\sqrt{2}}\sigma_\varepsilon \frac{\phi\left(\frac{v(\bar{\omega}_t, q_t, d_t=1) - \bar{U}(t, q_t, \bar{\omega}_t)}{\beta^t \sqrt{2}\sigma}\right)}{1 - \Phi\left(\frac{v(\bar{\omega}_t, q_t, d_t=1) - \bar{U}(t, q_t, \bar{\omega}_t)}{\beta^t \sqrt{2}\sigma}\right)} \tag{28}
$$

With $D > 1$, expressions are equivalently defined but no longer have closed form solutions. In the estimation, we simulate to obtain expressions for these conditional expectations.

B.4 Extension to multidimensional q

Where $dim(q) > 2$, we require additional "preference shifters", z, to identify beliefs. This proof relies on a standard exclusion restriction: z can affect preferences but not beliefs. We work through the proof for the specific binary choice case ($D = 1$) but the same principles apply to the general $D > 1$ case. Here we concentrate on the extension we use in the empirical application – 3 dimensions of offer quality (H, L, and N meaning that no offer has been received) but the same principles follow for introducing a "medium" quality groom.

We amend part (3b) of Assumption A1 and Assumption A3 to be:

Assumption A1,3b' $q_t \in \{H, L, N\}$ is a stochastic state variable that is observable ex-post to researchers (i.e. the quality of a marriage offer);

Assumption A3' Individuals' have proper beliefs over the realization of the stochastic state variable $q_t \in \{L, H, N\}$, where $0 \leq \pi^q(\bar{\omega}_t) \leq 1$ is the belief that $q_t = q$ given deterministic state variables $\bar{\omega}_t$.

In this case, a description of the value of z must be included in both the ex-ante and the ex-post experiments. Using the same argument as Theorem 1, in this case the ex-post experiment identifies preferences over states including z. Identification of $\pi(\bar{\omega}_t)$ then relies on making use in variation in ex-ante choice probabilities at different values of z.

Consider the case (as in our empirical application), where $q \in \{H, L, N\}$ with H and L defined as before and N represents the state of the world where no marriage offer is received. Let $\pi^H(\bar{\omega}_t)$ and $\pi^L(\bar{\omega}_t)$ give the probability of a high and low quality marriage offer and $1 - \pi^H(\bar{\omega}_t) - \pi^L(\bar{\omega}_t) \geq 0$ give the probability of no offer. Imagine we have access to a preference shifter with at least two distinct values $z' \neq z''$ and $u(\bar{\omega}_t, z', d_t = 1) \neq u(\bar{\omega}_t, z'', d_t = 1)$.

In this case, Step 1 of the proof above proceeds exactly as above except now everything is conditional on the value of the instrument z. We thus recover σ_ε and $v(\bar{\omega}_t, q, d_t = 1, z)$ for all $\bar{\omega}_t, q, d_t, z$.[75] In Step 2, the recursive relationship is modified to include the three possibilities for future realizations of q and

[75] Again here we use the fact that $v(\bar{\omega}_t, q', d_t = 1, z) = v(\bar{\omega}_t, q'', d_t = 1, z)$ for all q, q'' and that $v(q_t, \bar{\omega}_t, d_t = 0, z) \equiv \bar{U}(t, q_t, \bar{\omega}_t, z)$

the conditionality on z. In the case with binary choices, we have for both $z = z'$ and $z = z''$:

$$
\begin{aligned}
&v(\bar{\omega}_t, q_t, d_t = 1, z) \\
&\quad = \underbrace{\pi^L(\bar{\omega}_t)\overbrace{p^A(d_{t+1} = 0|\bar{\omega}_{t+1}, L)\left[v(\bar{\omega}_{t+1}, L, d_{t+1} = 0, z) + \beta\mathbb{E}\left(\varepsilon_t(0)|\bar{\omega}_{t+1}, L, d^*_{t+1} = 0, z\right)\right]}^{a(\bar{\omega}_{t+1}, z)}}_{\text{low offer, accept}} \\
&\qquad + \underbrace{\pi^H(\bar{\omega}_t)\overbrace{p^A(d_{t+1} = 0|\bar{\omega}_{t+1}, H)\left[v(\bar{\omega}_{t+1}, H, d_{t+1} = 0, z) + \beta\mathbb{E}\left(\varepsilon_t(0)|\bar{\omega}_{t+1}, H, d^*_{t+1} = 0, z\right)\right]}^{b(\bar{\omega}_{t+1}, z)}}_{\text{high offer, accept}} \\
&\qquad + \underbrace{\pi^L(\bar{\omega}_t)\overbrace{p^A(d_{t+1} = 1|\bar{\omega}_{t+1}, L)\left[v(\bar{\omega}_{t+1}, L, d_{t+1} = 1, z) + \beta\mathbb{E}\left(\varepsilon_t(1)|\bar{\omega}_{t+1}, L, d^*_{t+1} = 1, z\right)\right]}^{c(\bar{\omega}_{t+1}, z)}}_{\text{low offer, reject}} \\
&\qquad + \underbrace{\pi^H(\bar{\omega}_t)\overbrace{p^A(d_{t+1} = 1|\bar{\omega}_{t+1}, H)\left[v(\bar{\omega}_{t+1}, H, d_{t+1} = 1, z) + \beta\mathbb{E}\left(\varepsilon_t(1)|\bar{\omega}_{t+1}, H, d^*_{t+1} = 1, z\right)\right]}^{d(\bar{\omega}_{t+1}, z)}}_{\text{high offer, reject}} \\
&\qquad + \underbrace{(1 - \pi^H(\bar{\omega}_t) - \pi^L(\bar{\omega}_t))\overbrace{v(\bar{\omega}_{t+1}, N, d_{t+1} = 1, z)}^{e(\bar{\omega}_{t+1}, z)}}_{\text{no offer}}
\end{aligned} \tag{29}
$$

Stacking over two distinct $z' \neq z''$ gives a set of equations that can be solved for $\pi^H(\bar{\omega}_t)$ and $\pi^L(\bar{\omega}_t)$ whenever the instrument shifts utility profiles:

$$
\begin{aligned}
&\begin{bmatrix} v(\bar{\omega}_t, q_t, d_t = 1, z') - e(\bar{\omega}_{t+1}, z') \\ v(\bar{\omega}_t, q_t, d_t = 1, z'') - e(\bar{\omega}_{t+1}, z'') \end{bmatrix} \\
&= \begin{bmatrix} a(\bar{\omega}_t, z') + c(\bar{\omega}_t, z') - e(\bar{\omega}_t, z') & b(\bar{\omega}_t, z') + d(\bar{\omega}_t, z') - e(\bar{\omega}_t, z') \\ a(\bar{\omega}_t, z'') + c(\bar{\omega}_t, z'') - e(\bar{\omega}_t, z'') & b(\bar{\omega}_t, z'') + d(\bar{\omega}_t, z'') - e(\bar{\omega}_t, z'') \end{bmatrix} \begin{bmatrix} \pi^L(\bar{\omega}_t) \\ \pi^H(\bar{\omega}_t) \end{bmatrix}
\end{aligned} \tag{30}
$$

This system is invertible when the following condition holds:

$$
\begin{aligned}
&[a(\bar{\omega}_t, z') + c(\bar{\omega}_t, z') - e(\bar{\omega}_t, z')]\,[b(\bar{\omega}_t, z'') + d(\bar{\omega}_t, z'') - e(\bar{\omega}_t, z'')] \\
&\neq [a(\bar{\omega}_t, z'') + c(\bar{\omega}_t, z'') - e(\bar{\omega}_t, z'')]\,[b(\bar{\omega}_t, z') + d(\bar{\omega}_t, z') - e(\bar{\omega}_t, z')]
\end{aligned} \tag{31}
$$

This will hold (except for knife-edge cases) when $U(\tau, q_\tau, \bar{\omega}_\tau, z', \boldsymbol{\varepsilon}) \neq U(\tau, q_\tau, \bar{\omega}_\tau, z'', \boldsymbol{\varepsilon})$.

C Practical Design & Analysis Considerations

This Appendix provides some practical design considerations, and suggestions for robustness checks to perform in analysis, for researchers looking to implement our approach. It is intended as a complement to Section 2 of the main paper, which provides the formal assumptions underlying our approach, and our demonstration simulation and estimation code.[76] This list is not meant to be comprehensive but rather a useful starting point for researchers hoping to tailor a revealed belief measurement approach to their own context.

C.1 Experiment Design

1. **Decision Making Model.** The first thing carefully to consider is whether the decision making process under study can be expressed as a model that is consistent with assumptions A1, A2 and A3. The underlying model must be a finite-horizon, dynamic discrete choice model with features described in A1, with preferences described in A2 and beliefs described in A3. Ensuring that the underlying model fits these assumptions will involve considering: What should the relevant time period be? Is there a terminal action at every decision node? Can preferences be written as an additive sum of utility of the observed path taken through the model and the discounted sum of idiosyncratic preference shocks?

2. **Ex-Post Experiment Design.** This requires designing a discrete choice experiment that asks respondents to make binary choices between all paths through the model that end in a terminal action. In the notation of our paper, we refer to these paths as $\bar{\Psi}_j$. In some cases where researchers want to allow that no terminal action is chosen before the last period (e.g. in our case that the daughter remains unmarried in the last period), it may be useful to also include paths that end in the terminal period without the terminal action is being chosen (e.g. in our case, paths where the daughter is age 23, with education E and still unmarried).

3. **Ex-Ante Experiment Design.** This requires designing a discrete choice experiment that asks respondents to make discrete choices corresponding to choice at given decision nodes of the model. These experiments will specify the time-invariant and deterministic state variables that define the decision node $\bar{\omega}_t$ (in our case, age, education etc), the characteristics feeding into the stochastic variable (in our case characteristics of the groom and match). Respondents should be asked to choose between the terminal action and all other actions that are available at that node.

4. **Visual Aid.** Depending on the complexity of the choice problem and the number of characteristics that researchers want to specify in the vignettes, a visual aid may be appropriate to reduce

[76]The demonstration code is hosted at https://github.com/revealedbelief/revealedbelief

the cognitive load on respondents.

5. **Scenario Heterogeneity.** Researchers should consider which elements of heterogeneity they want to build into the model and whether these should enter preferences, beliefs or both. Characteristics driving heterogeneity in preferences should be included in both the ex-post and ex-ante experiments whereas characteristics driving heterogeneity in beliefs should only be reflected in the ex-ante experiments. In our case, this involved building in heterogeneity in a daughter's behavior (whether or not they complied with gender norms), in parental wealth, in the need for domestic help at home and in the daughter's attitude towards school.

6. **Anchoring Scenario Characteristics.** Where possible, researchers will find it useful to anchor the definitions and descriptions of scenario characteristics to those used in observational data from the same (or similar) samples. Having exact analogues of these characteristics in the observational data will help test for consistency between the experimental results and patterns in the observational data. For instance, in our case, we anchored descriptors of the wealth of households to quintiles of the observed asset ownership of households in our sample.

7. **Sampling Choice Scenarios.** The sampling of experimental choice scenarios will affect the precision of estimates. In general, how researchers approach this will depend on the strength of their priors about preferences and beliefs. If a researcher has a strong prior about the direction and magnitude of preference and belief parameters, it can be useful to weight the sampling of choice scenarios to those where the researcher does not predict a huge difference in how the different options will be valued (i.e. where there is not one clearly preferred option). If they have more diffuse priors they may want to simply ensure a good spread of scenarios across the entire state space; heavily weighing towards choice scenarios where the researcher *expects* respondents to be close to indifference may risk harming power if these priors are wrong. If researchers are hoping identify individual-level heterogeneity, or the parameters characterizing the distribution of such heterogeneity, they could consider building an algorithm to optimally select the next choice scenario to maximise new information. This could involve extending the algorithms of Drake et al. (2022) for preference estimation to the revealed belief case. Researchers can simulate the power and precision of the estimators for a given model set-up and given parameterizations of preferences and beliefs by adapting the demonstration code provided with this paper.[77]

[77] See https://github.com/revealedbelief/revealedbelief

8. **Tests for the Importance of Vignette Salience.** Researchers may wish to test whether respondents' choice behavior depends on the salience of a choice scenario, i.e. how close it is to their own situation. In this case, researchers could consider using scenario characteristics based on a respondent's own situation in one round of the experiment. Researchers should randomly vary which round is the "salient" round.

9. **Sample Size (Respondents and Rounds).** The sample size will clearly affect the power and precision of estimates. Researchers can explore the implications of varying the number of respondents and number of rounds using simulations described in the "sampling choice scenarios" point above. In general, having many rounds per respondent may permit allowing for some individual-level heterogeneity in preferences and/or beliefs. Even if this is not precisely identified at the individual level, its distribution across the sample may be.

10. **Piloting.** We strongly recommend multiple rounds of thorough piloting of these instruments. During this process researchers should check whether the most critical scenario characteristics (as perceived by respondents) are indeed adequately captured in the choice experiments. Open-ended feedback from respondents should be used following trials of the instruments to assess whether respondents interpreted the scenarios in the manner intended, where any misunderstandings arose, whether respondents paid attention to all aspects of the choice scenarios and whether respondents' reported reasoning in making choices was consistent with the underlying structure of the model. We would recommend documenting this stage carefully, to allow researchers to report on instrument development in any later published work.

11. **Measures of Complexity and Understanding Checks.** Researchers could consider following the choice scenarios with measures of how complex the respondent found the task and checks of understanding. These could take the form of a subjective enumerator assessment, a likert-scale based measure directly asking the respondent or a more formal measure of subjective complexity based on binary comparisons as described in Gabaix and Graeber (2023).

C.2 Analysis Robustness

1. **Functional Form.** In many applications, researchers will have to impose a specific functional form on preferences and beliefs for estimation. Researchers should conduct robustness checks to assess whether or not conclusions are sensitive to assumptions made in estimation. First, it

is useful to assess robustness to using different functional forms for both preferences and beliefs (we showed such checks in Tables II and A.4). Second, researchers should explore whether results are sensitive to varying the assumed discount factor (we showed this check in Table A.4). Third, researchers may want to explore the impacts of removing the possibility that respondents may be inattentive (we showed this check in Table A.4).

2. **Consistency of Taste Shock Scale.** The dynamic choice model underlying the revealed belief approach implies a specific relationship between the scale of the unobserved drivers of choice behavior across the ex-post and ex-ante experiments. In the ex-post experiment, the aggregated taste shocks comprise the discounted sum of per-period taste shocks. In our case with 10 periods and setting $\beta = 0.95$ this is $\nu_{i,j} \equiv \sum_{t=1}^{10} 0.95^t \varepsilon_{i,t,j}(d_{t,j})$.[78] This implies that $\sigma_\varepsilon = \sqrt{\sigma_\nu^2 / \sum_{t=1}^{10} 0.95^{2t}}$. Plugging in our estimate of $\hat{\sigma}_\nu = 1.202$ from the ex-post experiment of gives an implied σ_ε of 0.450 which is well within the confidence interval for our direct estimate of σ_ε from the ex-ante experiment (Table III).

3. **Patterns of Heterogeneity.** Researchers may have strong priors on the heterogeneity in preference and belief parameters based on observable characteristics of respondents. In this case, it is natural to test whether these priors are borne out in the empirical results. In cases (like our own) where respondents' are asked about how a fictional or typical decision maker would behave, we would typically have weaker priors on the relationship between a respondent's own characteristics and their choice behavior. In this instance, heterogeneity by observable characteristics could reflect different respondents having different reference points for the preferences and beliefs of a typical decision maker. Similarly, heterogeneity by proxies of respondents' own exposure to the type of situations being asked about (e.g. whether the respondent has gone through the marriage process for one of their own children) may be indicative of learning from experience and could suggest that respondents without such experience may have beliefs based on more limited information.

4. **Impact of Vignette Salience .** If researchers built in "salient" choice scenarios into their design then a check of whether choices vary with vignette salience is a natural check of whether respondents were able to position themselves in and consider choice scenarios that were more removed from their own situation. Even when researchers did not explicitly build in such salient scenarios,

[78] See Section 2.3.

randomly occurring variation in scenario characteristics leading to some choice scenarios being more similar to a respondents own situation than others, could also be used for this purpose.

5. **Inattention.** We follow Mas and Pallais (2017) in suggesting that researchers estimate the inattention rate directly from the discrete choice experiments. Researchers can then check whether the inattention rate estimated appears plausible and in line with other studies.[79] To assess whether excess cognitive load is a major concern, researchers may wish to check whether the inattention rate varies with respondent characteristics that we might expect to be correlated with cognitive load (such as education) or by round of the experiment to assess whether growing fatigue caused some respondents to become inattentive.

6. **Evidence of Cognitive Overload.** Estimated inattention rate and heterogeneity therein will be informative about whether cognitive overload was a problem. In addition, supplementary measures of cognitive load, complexity and fatigue (either self reported by the respondent or assessed by the interviewer) can be analyzed and correlated with estimated preferences and beliefs to assess whether such problems may have led to attenuation of the patterns found.

7. **Consistency with Other Measurements.** In many cases, it may be useful to check the consistency of the preference and belief estimates with estimates of other objects from the same sample. Exactly what these other objects are and how they are best measured will of course depend on the exact application. In our marriage market context, we collected simpler measures of stated expectations about likely marriage offers at different points in time and preferences on the groom's side of the marriage market as checks of internal consistency. These are described in Section 6.4. Whether these internal consistency checks are qualitative or quantitative in nature will depend on the nature of the measures used and the appropriateness of any further assumptions that would need to be made in order to make the measures quantitatively comparable. For instance, in our case, for our estimates of preferences on the groom's side of the marriage market to be comparable with beliefs on the bride's side, we would have had to make very strong assumptions about the nature of frictions in the marital search process.

8. **Consistency with Observational Data.** If observational data on choice behaviour with scenario characteristics is available, one can check how the trajectories implied by the estimated preference

[79] For instance, Mas and Pallais (2017) estimate an inattention rate of 30.16% when estimating the rate internally using maximum likelihood estimation.

and belief parameters compare to the patterns in observational data. This is most useful when the observational data is from the same sample of respondents but a comparison to similar samples could be used when this is not possible. In Section 6.4, we compared the implied vs. predicted fraction of daughters who were in school and married at each age. Where researchers have built in heterogeneity into the model (in preferences, beliefs, or both) along dimensions that have analogues in the observational data, the magnitude of this heterogeneity can also be compared between the estimated predictions and the observed patterns.

While we think of this exercise as highly informative about the plausibility of the experimental estimates, we note that researchers shouldn't necessarily expect observed and predicted trajectories to be completely aligned. The predicted trajectories will capture the *expected* trajectories based upon respondents' beliefs at the time when the data was collected. Given that the observed trajectories necessarily result from choices made over a prolonged period of time, any unanticipated aggregated shocks to the distribution of the stochastic variable of interest may cause a divergence that is perfectly compatible with economic theory.

C.3 Possible Extensions Checklist

1. **Preference Shifting Instruments.** In Appendix B.4, we show how the use of instruments that shift preferences but are excluded from beliefs allows for the identification of beliefs when the stochastic state variable q has more than two dimensions. i.e. rather than impose that q has two dimensions (e.g. low versus high), one might want to allow for q to be high-dimensional (e.g. low versus medium versus high). In our application, we use whether or not the fictional daughter "likes school" as an instrument to allow us to identify the beliefs over the probability of not receiving a marriage offer at all in a given period.

2. **Own Choice *vs.* Choice of a Fictional or Typical Decision Maker.** In our application, we always asked respondents about the choices they thought that a fictional couple we described to them would make for their fictional daughter. We made this choice to bypass concerns about social desirability bias and to bypass concerns about unobserved heterogeneity driving choice behavior. In other applications, researchers may want to ask respondents directly about what they themselves would choose (in response to different hypothetical scenarios) which would allow researchers to interpret estimates, and heterogeneity by observables, as indicative of respondents own preferences and beliefs rather than their beliefs about typical preferences and beliefs.

Researchers could usefully compare the impact of these two framings by randomizing whether respondents are asked about their own choice vs. the choice of a fictional decision maker and comparing the resulting estimates.

3. **Aggregation of Preferences & Beliefs into Collective Choice.** In our application, were interested in understanding the *collective* choice of married couples over their daughter's education and marriage. We thus did not make a distinction between the preferences of beliefs of the husband or wife or seek to understand how these might be combined to form collective preferences and beliefs. In other settings, researchers might find it productive to expand our approach to assess separately the preferences and beliefs of different individuals involved in collective choice (perhaps in addition to also measuring collective preferences and beliefs). Researchers could consider, for instance, asking the same respondents about their perceptions of what both wives and husbands would separately choose in addition to what their collective choice would be. Researchers could also compare the choices of male and female respondents to such questions to assess whether they have different insights or information about preferences of different groups.

D Summaries of Focus Group Discussions

Click here to view the translated summaries of the three focus groups carried out with caregivers of adolescent girls as part of formative research.