Appendix: For Online Publication Only
Who You Gonna Call? Gender Inequality in External Demands for Parental Involvement

Kristy Buzard[1], Laura Gee[2], and Olga Stoddard[3]

[1]Syracuse University
[2]Tufts University
[3]Brigham Young University

August 28, 2025

Contents

A Appendix Tables

Table A.1: **Multinomial Logit Models of Effect of Treatments**

	(1)	(2)
Outcome: Female Call		
High Male (Hm)	-0.82***	-0.55***
	(0.07)	(0.07)
Low Female (Lf)	-0.23***	-0.04
	(0.06)	(0.06)
Low Male (Lm)	0.22***	0.16**
	(0.05)	(0.06)
High Female (Hf)	0.48***	0.67***
	(0.05)	(0.05)
Outcome: Male Call		
High Male (Hm)	0.62***	0.60***
	(0.06)	(0.06)
Low Female (Lf)	0.25***	0.14*
	(0.07)	(0.07)
Low Male (Lm)	-0.39***	-0.23**
	(0.08)	(0.07)
High Female (Hf)	-1.32***	-0.84***
	(0.10)	(0.08)
Observations	30471	30320

Notes: This table presents the results of a multinomial logit model using a model like the one in Equation 1. The outcome variable takes three values: no call, call female, or call male. In this table we present the results with a base case of no call. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent. The outcomes from this table are represented visually in Figure B.1. + $p < 0.10$ * $p < 0.05$ ** $p < 0.010$ *** $p < 0.001$

Table A.2: **Selection Into NoCall by Observable Variables of Schools By Variation**

	(1) Baseline	(2) Equal Decision	(3) Full-Time	(4) Pay
Other Schools	-0.02	-0.06**	-0.05	0.06
	(0.03)	(0.02)	(0.04)	(0.06)
Middle	-0.03+	-0.02	-0.04	-0.02
	(0.02)	(0.01)	(0.03)	(0.02)
High	0.02+	-0.01	0.02	0.00
	(0.01)	(0.01)	(0.02)	(0.02)
Decison-Maker Female	0.01	0.02+	0.01	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)
Public (non-Charter)	0.09***	0.07***	0.04+	0.14***
	(0.02)	(0.01)	(0.02)	(0.03)
Public (Charter)	0.11***	0.06**	0.01	0.04
	(0.03)	(0.02)	(0.04)	(0.05)
High Male (Hm)	0.02	-0.04	-0.12	0.14+
	(0.05)	(0.04)	(0.08)	(0.08)
Low Female (Lf)	0.02	-0.06	0.02	0.08
	(0.05)	(0.04)	(0.07)	(0.08)
Low Male (Lm)	0.02	0.03	0.02	0.08
	(0.05)	(0.03)	(0.07)	(0.08)
High Female (Hf)	0.02	-0.05	-0.15*	-0.02
	(0.05)	(0.04)	(0.07)	(0.10)
Other Schools * MaleHigh	-0.02	0.01	0.03	-0.15*
	(0.04)	(0.04)	(0.07)	(0.07)
Other Schools * FemaleLow	0.01	0.04	-0.08	-0.04
	(0.04)	(0.04)	(0.06)	(0.08)
Other Schools * MaleLow	-0.01	-0.03	-0.07	-0.06
	(0.04)	(0.03)	(0.06)	(0.08)
Other Schools * FemaleHigh	-0.04	0.01	0.06	-0.05
	(0.04)	(0.04)	(0.07)	(0.08)
Middle * MaleHigh	-0.01	-0.01	0.06	0.02
	(0.02)	(0.02)	(0.04)	(0.03)
Middle * FemaleLow	0.01	0.02	0.06	-0.02
	(0.02)	(0.02)	(0.04)	(0.03)
Middle * MaleLow	0.03	-0.01	0.03	-0.02
	(0.02)	(0.02)	(0.04)	(0.03)
Middle * FemaleHigh	0.03	-0.01	-0.02	-0.01
	(0.02)	(0.02)	(0.03)	(0.04)
High * MaleHigh	-0.01	0.01	0.01	0.01
	(0.02)	(0.02)	(0.03)	(0.03)
High * FemaleLow	-0.02	0.02	0.02	-0.03
	(0.02)	(0.02)	(0.03)	(0.03)
High * MaleLow	-0.02	0.02	0.02	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)
High * FemaleHigh	-0.02	0.00	-0.02	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)
Decison-Maker Female * MaleHigh	0.01	0.01	-0.00	0.03
	(0.02)	(0.02)	(0.03)	(0.03)
Decision-Maker Female * FemaleLow	-0.01	-0.03+	-0.01	0.04
	(0.02)	(0.02)	(0.03)	(0.03)
Decison-Maker Female * MaleLow	-0.01	-0.01	-0.03	0.01
	(0.02)	(0.01)	(0.03)	(0.02)
Decison-Maker Female * FemaleHigh	0.01	-0.01	0.02	0.03
	(0.02)	(0.02)	(0.03)	(0.03)
Public (non-Charter) * MaleHigh	-0.00	-0.00	0.07+	-0.04
	(0.02)	(0.02)	(0.04)	(0.04)
Public (non-Charter) * FemaleLow	-0.00	0.02	0.03	-0.06
	(0.02)	(0.02)	(0.04)	(0.04)
Public (non-Charter) * MaleLow	-0.01	-0.00	0.05	-0.04
	(0.03)	(0.02)	(0.04)	(0.04)
Public (non-Charter) * FemaleHigh	0.01	0.01	0.07*	0.05
	(0.02)	(0.02)	(0.03)	(0.05)
Public (Charter) * MaleHigh	-0.03	0.00	0.09	0.07
	(0.04)	(0.04)	(0.07)	(0.07)
Public (Charter) * FemaleLow	-0.08*	0.04	-0.00	-0.05
	(0.04)	(0.04)	(0.07)	(0.07)
Public (Charter) * MaleLow	-0.03	0.02	0.08	0.06
	(0.04)	(0.03)	(0.06)	(0.06)
Public (Charter) * FemaleHigh	-0.02	0.03	0.11+	0.12+
	(0.04)	(0.04)	(0.06)	(0.07)
Constant	0.73***	0.81***	0.84***	0.68***
	(0.03)	(0.03)	(0.05)	(0.07)
R^2	0.01	0.01	0.01	0.02
Observations	30471	30320	9472	9808

Notes: "Other schools" are K-12 or pre-schools. Decision-Maker Female is whether the decision-maker (the principal) has a first name that is female. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent) + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$..

Table A.3: **Multinomial Logit Models For Theory Model**

	(1) No Call Base	(2) RegDeepParam2
Panel A: Outcome Female Call (vs. No Call)		
Any Signal About Male	-0.23***	
	(0.07)	
x_M (Male Signal Pos/Neg)	-0.44***	
	(0.04)	
Any Signal About Female	0.20**	
	(0.07)	
x_F (Female Signal Pos/Neg)	0.57***	
	(0.04)	
reply-to-sender*MaleHigh	0.70***	0.70***
	(0.07)	(0.07)
reply-to-sender*MaleLow	0.35***	0.35***
	(0.04)	(0.04)
reply-to-sender*FemaleHigh	0.20***	0.20***
	(0.03)	(0.03)
reply-to-sender*FemaleLow	1.01***	1.01***
	(0.07)	(0.07)
reply-to-sender*NoSignal	0.80***	0.80***
	(0.06)	(0.06)
High Male (Hm)		-0.67***
		(0.09)
High Female (Hf)		0.77***
		(0.06)
Low Male (Lm)		0.21**
		(0.07)
Low Female (Lf)		-0.36***
		(0.09)
Constant	-2.17***	-2.17***
	(0.06)	(0.06)
Panel B: Outcome Male Call (vs. No Call)		
Any Signal About Male	0.75***	
	(0.14)	
x_M (Male Signal Pos/Neg)	0.79***	
	(0.06)	
Any Signal About Female	0.14	
	(0.15)	
x_F (Female Signal Pos/Neg)	-0.52***	
	(0.07)	
reply-to-sender*MaleHigh	-0.34***	-0.34***
	(0.04)	(0.04)
reply-to-sender*MaleLow	-1.45***	-1.45***
	(0.10)	(0.10)
reply-to-sender*FemaleHigh	-1.12***	-1.12***
	(0.12)	(0.12)
reply-to-sender*FemaleLow	-1.05***	-1.05***
	(0.07)	(0.07)
reply-to-sender*NoSignal	-1.72***	-1.72***
	(0.13)	(0.13)
High Male (Hm)		1.54***
		(0.14)
High Female (Hf)		-0.38*
		(0.18)
Low Male (Lm)		-0.03
		(0.17)
Low Female (Lf)		0.66***
		(0.15)
Constant	-3.32***	-3.32***
	(0.13)	(0.13)
Observations	29363	29363

Notes: Column 1 of this table presents the results of a multinomial logit model using a model like the one in Equation 1. The outcome variable takes three values: no call, call female, or call male. The right-hand side variables are "Any Signal About Male" which takes the value 1 if a message was sent with a signal about the male parent (MaleHigh, MaleLow) and zero otherwise. "Any Signal About Female" takes the value 1 if a message with a signal about the female parent was sent (FemaleHigh, FemaleLow) and zero otherwise. The variable x_M (Male Signal Pos/Neg) takes the value 1 if the MaleHigh message was sent, and -1 if the MaleLow message, 0 otherwise; x_F is defined analogously for messages about female parents. Column 2 of this table presents results that are discussed in Appendix Section H.F. In both modelts there are a series of variables that control for the gender of the sender of the email (male vs. female parent) interacted with the signals about each parent's availability. The variable "reply-to-sender" takes the value 1 if the sender of the email is female, and -1 if the sender of the email is male (recall we always send an email from one parent and CC the other parent). The variables which capture which of our five messages were sent (MaleHigh, MaleLow, FemaleHigh, FemaleLow and NoSignal) are interacted with "reply-to-sender." The right-hand side variables are discussed in Section H. In this table we present the results with a base case of no call. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent). + $p <$ 0.10 * $p < 0.05$ ** $p < 0.010$ *** $p < 0.001$

Table A.4: **Reply-to-Sender: Likelihood of Call To Mother By Whether Mother Sent The Email**

	(1) Called Female | Call	(2) Called Female | Call
High Male (Hm)	0.12***	0.11***
	(0.01)	(0.01)
Low Female (Lf)	0.08***	0.13***
	(0.01)	(0.02)
Baseline	0.20***	0.18***
	(0.02)	(0.02)
Low Male (Lm)	0.48***	0.40***
	(0.02)	(0.02)
High Female (Hf)	0.83***	0.73***
	(0.01)	(0.02)
FemEmailMaleHigh	0.27***	0.39***
	(0.02)	(0.03)
FemEmailFemaleLow	0.78***	0.77***
	(0.02)	(0.02)
FemEmailBaseline	0.77***	0.77***
	(0.02)	(0.02)
FemEmailMaleLow	0.48***	0.56***
	(0.02)	(0.02)
FemEmailFemaleHigh	0.14***	0.25***
	(0.02)	(0.02)
R^2	0.80	0.79
Observations	6319	5980

	(1) Called Male | Call	(2) Called Male | Call
High Male (Hm)	0.61***	0.51***
	(0.02)	(0.02)
Low Female (Lf)	0.14***	0.10***
	(0.01)	(0.01)
Baseline	0.02***	0.05***
	(0.01)	(0.01)
Low Male (Lm)	0.04***	0.04***
	(0.01)	(0.01)
High Female (Hf)	0.03***	0.03***
	(0.01)	(0.01)
MalEmailMaleHigh	0.27***	0.39***
	(0.02)	(0.03)
MalEmailFemaleLow	0.78***	0.77***
	(0.02)	(0.02)
MalEmailMaleLow	0.48***	0.56***
	(0.02)	(0.02)
MalEmailFemaleHigh	0.14***	0.25***
	(0.02)	(0.02)
MalEmailBaseline	0.77***	0.77***
	(0.02)	(0.02)
R^2	0.71	0.70
Observations	6319	5980

Notes: In this table we regress the likelihood the mother was called (conditional on a call being made) on availability message sent (HighMale, LowMale, NoSignal, HighFemale, LowFemale) and the interaction of the message with whether the email was sent from the mother's email (CCing the father) and listing the mother's name and phone number first. We would obtain the same estimates of the interaction terms were we to run a regression of likelihood the father was called (conditional on a call being made) on the same set of right-hand-side variables. This allows us to break down the total calls to mothers into those made because she sent the email and residual reasons.

Table A.5: **More vs. Less Traditional Gender Norms Summary Statistics No Signal Message in Baseline Variation**

	(1) Non Religious School	(2) Religious School	(3) Low Repub. County	(4) High Repub. County	(5) Small Wage Gap County	(6) Large Wage Gap County	(7) Less Rural County	(8) More Rural County	(9) Less Religious County	(10) More Religious County	(11) Less Sexist State	(12) More Sexist State
Called Female	0.11	0.21	0.08	0.13	0.08	0.15	0.12	0.14	0.13	0.13	0.13	0.12
Called Male	0.08	0.12	0.07	0.09	0.04	0.12	0.08	0.09	0.09	0.06	0.07	0.07
No Call	0.81	0.67	0.85	0.78	0.88	0.74	0.80	0.77	0.78	0.81	0.79	0.81
Called Female | Call	0.58	0.63	0.51	0.58	0.70	0.56	0.59	0.62	0.58	0.69	0.64	0.61
Called Male | Call	0.42	0.37	0.49	0.42	0.30	0.44	0.41	0.38	0.42	0.31	0.36	0.39
Observations	4755	528	631	580	531	593	4445	1164	609	555	485	607

Notes: Religious school means the school is identified by our schools database as a religious school, while Non-Religious schools include public schools (non-charter) and private schools (non-religious). Low Republican means the school is located in a county at the 10th percentile or below of Republican vote share in the 2016 presidential election, while High Republican is at the 90th percentile or above. Small Wage Gap means the school is located in a county at the 10th percentile or below of the ratio between male-female median wages, while Large Wage Gap is at the 90th percentile or above. More Rural county means fewer than 250,000 population, while Less Rural is above that. Less Religious county is a county at the 10th percentile or lower for religious adherence, while More Religious county is above the 90th percentile as measure by the Association of Statisticians of American Religious Bodies (https://www.thearda.com/us-religion/sources-for-religious-congregations-membership-data#QR). Less Sexist State means the school is located in a state at the 10th percentile or below of the sexism index created by questions from the General Social Survey, while High Sexist State is at the 90th percentile or above (Kerwin et al., 2022). Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

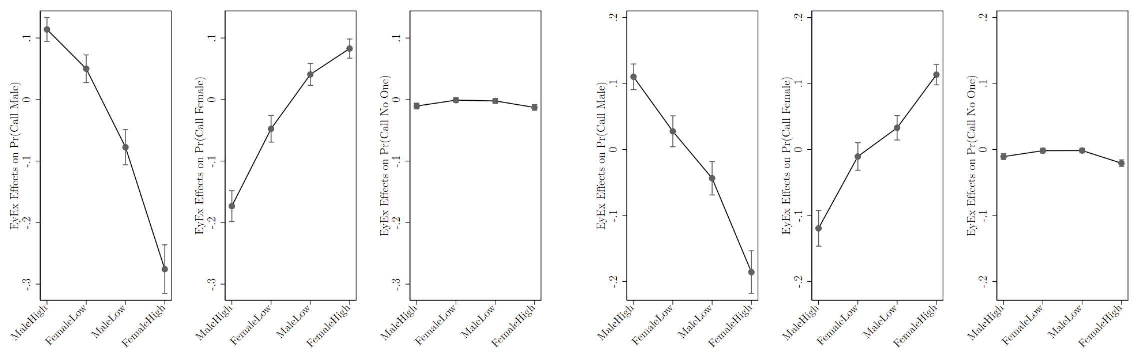
Table A.6: **More vs. Less Traditional Gender Norms Summary Statistics No Signal Message in Equal Variation**

	(1) Non Religious School	(2) Religious School	(3) Low Repub. County	(4) High Repub. County	(5) Small Wage Gap County	(6) Large Wage Gap County	(7) Less Rural County	(8) More Rural County	(9) Less Religious County	(10) More Religious County	(11) Less Sexist State	(12) More Sexist State
Called Female	0.09	0.17	0.08	0.11	0.08	0.12	0.10	0.11	0.11	0.11	0.09	0.09
Called Male	0.08	0.12	0.04	0.09	0.05	0.08	0.08	0.10	0.09	0.09	0.07	0.08
No Call	0.83	0.72	0.88	0.79	0.87	0.80	0.82	0.80	0.80	0.80	0.84	0.83
Called Female | Call	0.55	0.59	0.64	0.55	0.63	0.58	0.58	0.53	0.54	0.53	0.54	0.55
Called Male | Call	0.45	0.41	0.36	0.45	0.37	0.42	0.42	0.47	0.46	0.47	0.46	0.45
Observations	5367	825	849	655	720	631	5210	1350	631	655	607	697

Notes: Variables are defined as in Table A.5.

B Appendix Figures

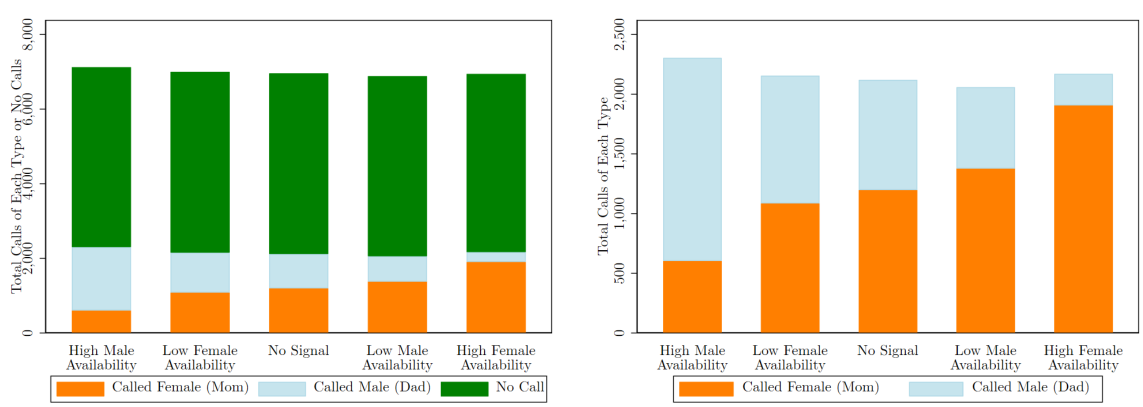
Figure B.1: Effects by Treatment



(a) Baseline Variation (b) Equal Decision Variation

Notes: In this figure we show the results from a mutltinomial logit model using a model like Equation 1 which is detailed fully in Table A.1. This figure shows the marginal effects elasticities. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

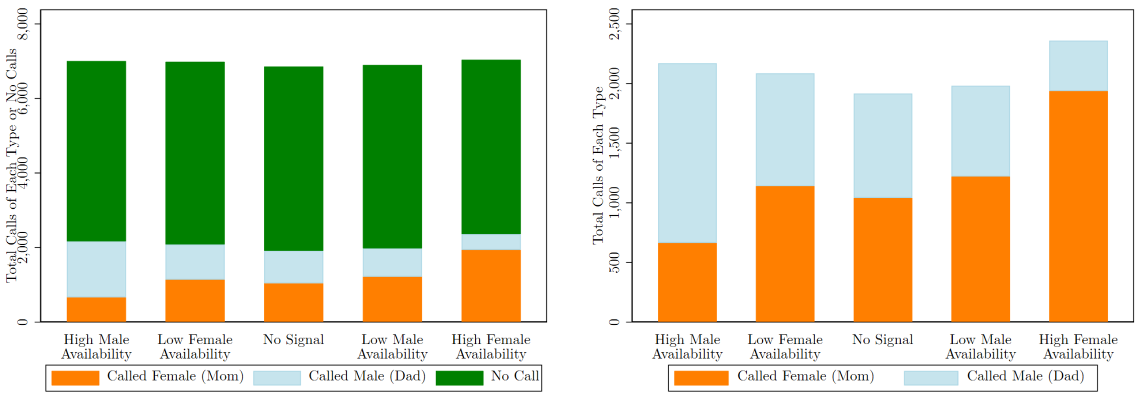
Figure B.2: Outcomes by Treatment in Baseline Variation for Multiple Calls



(a) All Outcomes (b) Outcomes Conditional on Calling

Notes: In this figure we show the total number of no calls, calls the female parent (mom) or calls to the male parent (dad) by the message sent to the decision-maker in our Baseline variation (see Figure 3 for proportions by only the first call or no call). Panel (a) represents three outcomes from 30, 471 decision-makers, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6, 382$). If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents. Two-way t-tests comparing No Call, Call Female, and Call Male are all statistically significant at the 5% level or below. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

Figure B.3: Outcomes by Treatment in Equal Decision Variation for Multiple Calls



(a) All Outcomes

(b) Outcomes Conditional on Calling

Notes: In this figure we show the total number of no calls, calls the female parent (mom) or calls to the male parent (dad) by the message sent to the decision-maker in our Baseline variation (see Figure 3 for proportions by only the first call or no call). Panel (a) represents three outcomes from 30,320 decision-makers, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6,046$). If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents. Two-way t-tests comparing No Call, Call Female, and Call Male are all statistically significant at the 5% level or below. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

C Balance Tables

See Tables F.1, and F.2 for balance in the other Variations of our experiment.

Table C.1: **Balance on Observable Attributes of Schools/Decision-makers by Treatment in Baseline Variation**

	(1) High Male	(2) Low Female	(3) Baseline	(4) Low Male	(5) High Female
Elementary	0.50	0.51	0.53	0.52	0.52
Middle	0.21	0.21	0.19	0.21	0.20
High	0.25	0.25	0.25	0.24	0.24
Decison-Maker Female	0.57	0.58	0.59	0.59	0.58
Public (Charter)	0.06	0.05	0.06	0.06	0.06
Public (non-Charter)	0.76	0.79	0.81	0.79	0.80
Private	0.18	0.16	0.13	0.15	0.14
Free Lunch	0.55	0.56	0.54	0.55	0.52
White	0.52	0.52	0.52	0.53	0.52
Black	0.14	0.15	0.14	0.14	0.15
Hispanic	0.23	0.23	0.23	0.23	0.23
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	7075	5931	5612	5700	6153

Notes: There is a small proportion of schools which are not elementary, middle or high schools (e.g. K–12 or preschools). The following variables are known only for non-private schools: FreeLunch, White, Black, Hispanic. DMFemale is whether the decision-maker (the principal) has a first name that is female. Observations are weighted so that 50% of emails are from a female parent and 50% from a male parent.

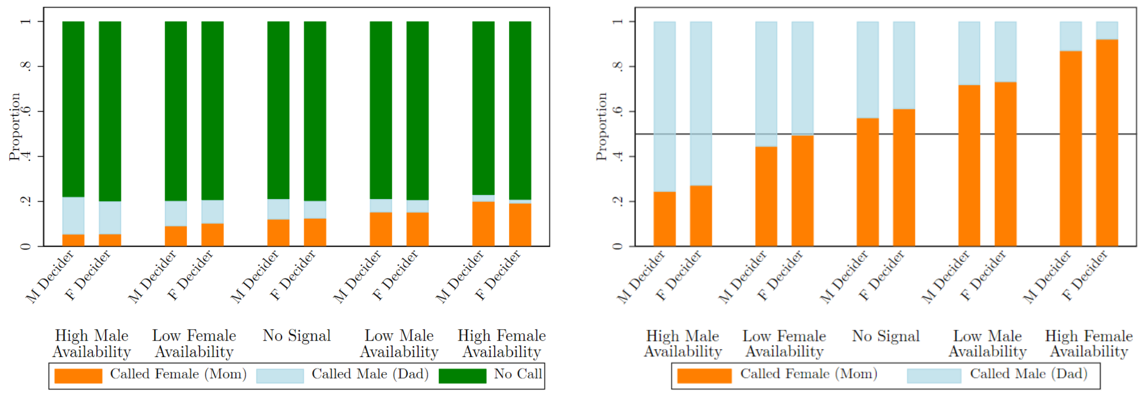
Table C.2: **Balance on Observable Attributes of Schools/Decision-Makers By Treatment In Equal Decision Variation**

	(1) High Male	(2) Low Female	(3) Baseline	(4) Low Male	(5) High Female
Elementary	0.52	0.52	0.51	0.50	0.50
Middle	0.21	0.21	0.21	0.22	0.22
High	0.25	0.24	0.24	0.25	0.25
Decison-Maker Female	0.58	0.58	0.57	0.58	0.57
Public (Charter)	0.06	0.05	0.06	0.06	0.05
Public (non-Charter)	0.80	0.80	0.77	0.76	0.76
Private	0.14	0.14	0.18	0.18	0.18
Free Lunch	0.55	0.52	0.55	0.55	0.57
White	0.52	0.53	0.52	0.52	0.52
Black	0.15	0.15	0.15	0.15	0.15
Hispanic	0.23	0.23	0.23	0.23	0.23
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	5170	5558	6569	6755	6268

Notes: Notes are the same as those in Table C.1.

D Decision-Maker Gender

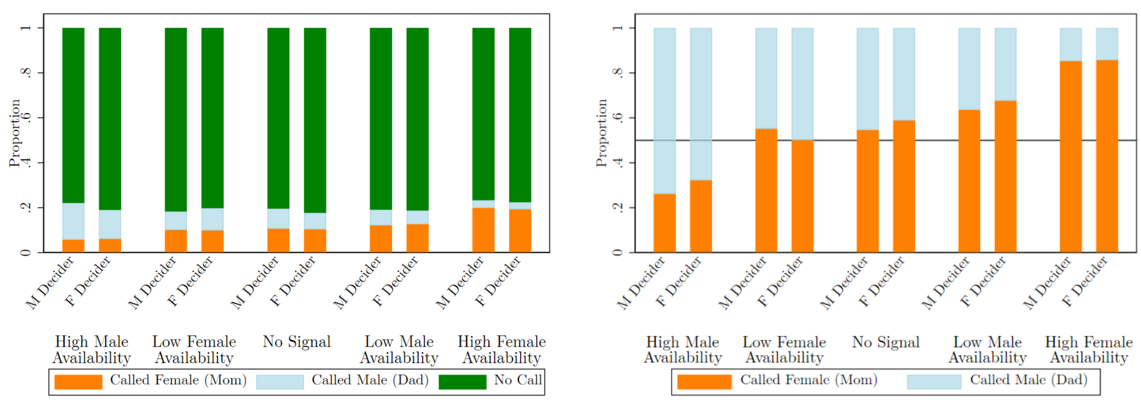
Figure D.1: Outcomes By Principal Gender in Baseline Variation



(a) All Outcomes (b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between Female and Male principals. We predict principal gender based on their name. In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision-maker in our Baseline Variation. "M Decider" denotes a male principal and "F Decider" denotes a female principal. Panel (a) represents three outcomes from 30, 471 decision-makers in Main, while panel (b) shows only the choices of those who made a phone call to at least one parent. In Panel B we show the breakdown for only those who called back.

Figure D.2: Outcomes By Principal Gender in Equal Decision Variation

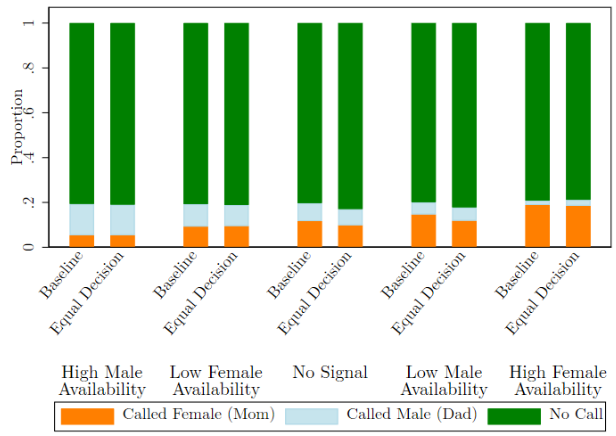


(a) All Outcomes (b) Outcomes Conditional On Calling

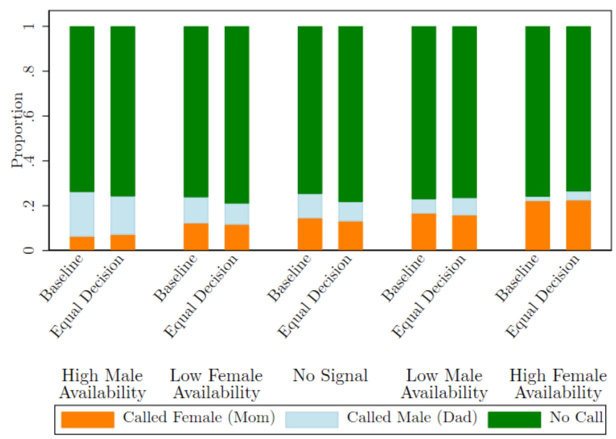
Notes: Notes are the same as in Figure D.1.

E By Grade-Level

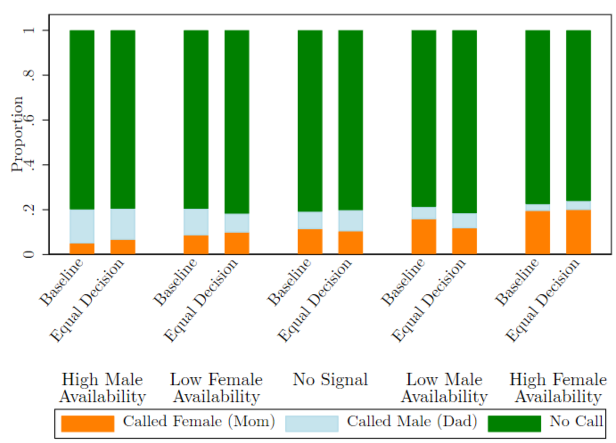
Figure E.1: Outcomes By Grade-Level



(a) Elementary



(b) Middle School



(c) High School

Notes: In this figure, we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision-maker in our Baseline and Equal Decision variations.

F Variations On Baseline Messages

Table F.1: **Balance on Observable Attributes of Schools/Decision-Makers By Treatment In Full Time Variation**

	(1) High Male	(2) Low Female	(3) Baseline	(4) Low Male	(5) High Female
Elementary	0.50	0.52	0.50	0.53	0.52
Middle	0.23	0.21	0.21	0.20	0.21
High	0.25	0.25	0.25	0.24	0.25
Decison-Maker Female	0.56	0.59	0.57	0.60	0.59
Public (Charter)	0.06	0.06	0.05	0.06	0.05
Public (non-Charter)	0.80	0.82	0.73	0.81	0.77
Private	0.14	0.12	0.22	0.13	0.18
Free Lunch	0.55	0.56	0.53	0.55	0.54
White	0.52	0.52	0.52	0.53	0.52
Black	0.15	0.15	0.14	0.15	0.14
Hispanic	0.23	0.23	0.24	0.22	0.24
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	1785	1478	1943	1776	2490

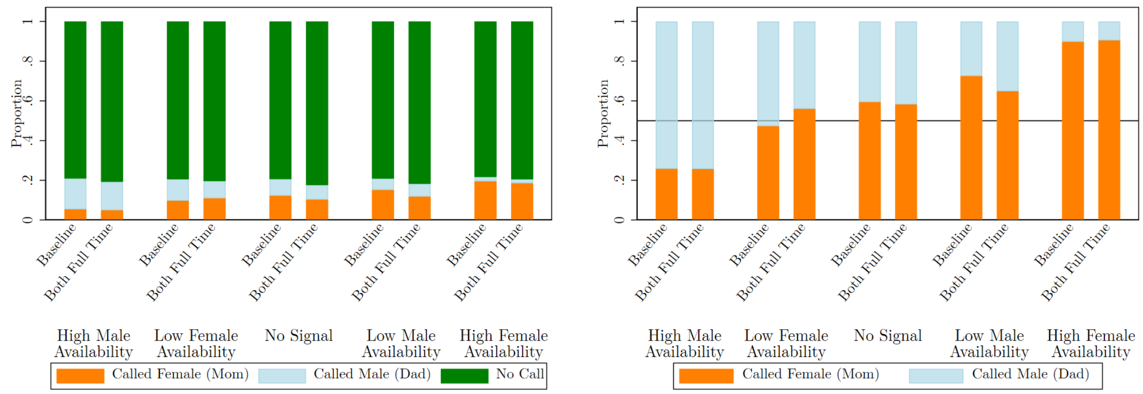
Notes: There is a small proportion of schools which are not Elementary, Middle or High Schools (e.g. K-12 or pre-schools). The following variables are only known for non-private schools: FreeLunch, White, Black, Hispanic. DMFemale is whether the decision-maker (the principal) has a first name that is female. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.

Table F.2: **Balance on Observable Attributes of Schools/Decision-Makers By Treatment In Payments Variation**

	(1) High Male	(2) Low Female	(3) Baseline	(4) Low Male	(5) High Female
Elementary	0.51	0.52	0.52	0.50	0.53
Middle	0.21	0.21	0.20	0.21	0.20
High	0.24	0.24	0.26	0.25	0.24
Decison-Maker Female	0.58	0.60	0.58	0.58	0.58
Public (Charter)	0.06	0.07	0.05	0.06	0.06
Public (non-Charter)	0.78	0.75	0.81	0.78	0.81
Private	0.17	0.18	0.14	0.16	0.12
Free Lunch	0.54	0.58	0.56	0.55	0.53
White	0.52	0.51	0.51	0.50	0.53
Black	0.15	0.15	0.15	0.15	0.15
Hispanic	0.23	0.23	0.23	0.25	0.22
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	2101	2153	1795	2333	1426

Notes: Notes are the same as Table F.1.

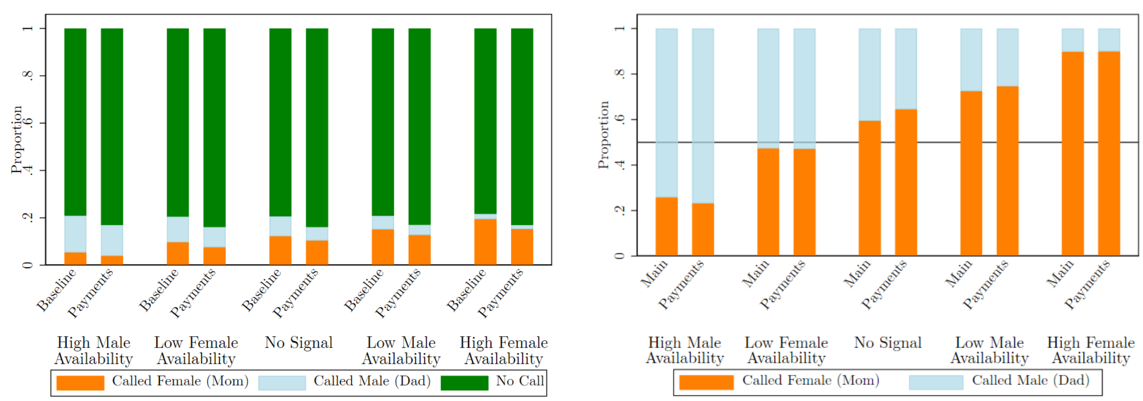
Figure F.1: Outcomes By Treatment "Baseline" vs. "Full Time" Variations



(a) All Outcomes (b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between our "Main" version of our emails and ones that have the addition of a sentence that states "We both work full-time." In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision-maker in our Baseline Variation. Panel (a) represents three outcomes from 30,471 decision-makers in Baseline and 9,472 in Full Time, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6382$ in Baseline and 1817 in Full Time).

Figure F.2: Outcomes By Treatment "Baseline" vs. "Payments" Variations



(a) All Outcomes (b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between our "Main" version of our emails and ones that have the addition of a clauses that states they are "especially interested in discussing school fees and other expenses." In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision-maker in our Baseline Variation. Panel (a) represents three outcomes from 30,471 decision-makers in Baseline and 9,808 in Full Time, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6382$ in Baseline and 1817 in Full Time). The patterns look similar if we restrict to private schools only.

G Example Emails Full Text

Figure G.1: Baseline: No Signal

School Inquiry

roy@miller-family.net <roy@miller-family.net>
To: laura.k.gee@gmail.com
Cc: erica@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

Roy(XXX)- XXX-8474 or Erica(XXX)- XXX-2761.

School Inquiry

erica@miller-family.net <erica@miller-family.net>
To: laura.k.gee@gmail.com
Cc: roy@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

Erica(XXX)- XXX-8505 or Roy(XXX)- XXX-8470.

Figure G.2: Baseline: High Female and Low Female Signal

School Inquiry

roy@miller-family.net <roy@miller-family.net>
To: laura.k.gee@gmail.com
Cc: erica@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

Erica has a lot of availability to chat, but you can call either me or Erica.

Roy (727) 855-3147 or Erica (727) 855-3137.

School Inquiry

erica@miller-family.net <erica@miller-family.net>
To: laura.k.gee@gmail.com
Cc: roy@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

I have limited availability to chat, but you can call either me or Roy.

Erica (727) 855-3125 or Roy (727) 855-3157.

Figure G.3: Baseline: High Male and Low Male Signal

School Inquiry

roy@miller-family.net <roy@miller-family.net>
To: laura.k.gee@gmail.com
Cc: erica@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

I have a lot of availability to chat, but you can call either me or Erica.

Roy (727) 855-3143 or Erica (727) 855-3100.

School Inquiry

erica@miller-family.net <erica@miller-family.net>
To: laura.k.gee@gmail.com
Cc: roy@miller-family.net

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to discuss?

Roy has limited availability to chat, but you can call either me or Roy.

Erica (727) 855-3121 or Roy (727) 855-3099.

H Theoretical Model

Our theoretical framework models how a decision-maker who interacts with a two-parent heterosexual household decides which person to call upon for a task. We built this model to inform the design of the experiment so that we can untangle the mechanisms that underlie any differential treatment of male versus female parents.

In our specific field experiment, the decision-maker is a school principal, and the task is a discussion about enrolling at the school. However, the model is flexible enough to be applied to different types of decision-makers (e.g., doctors, school teachers, sports coaches, organized religion leaders) and different kinds of tasks (e.g., picking up a sick child, communicating about health concerns, taking the team on an overnight trip). Furthermore, our model could apply outside of parenting tasks to study many types of demands on a two-person household (e.g., for elder care, home renovations, retirement planning) as long as the central elements are present: one decision-maker, a set of differentiated individuals to contact, and messages that inform key beliefs about the individuals to be contacted.

We lay out a simple economic structure in Section H.A to capture the decision-making behavior of school principals when contacting parents. In Section H.B, we describe the random utility model we have constructed to study this environment. We then explain in Section H.C how our experimental variation integrates with the random utility model. Section H.D shows how we use the model to identify and estimate its structural parameters, most notably the parameters for principals' beliefs and the other deterrents they face to calling parents. Section H.E outlines key testable hypotheses of interest. Appendix H contains additional model details as well as all proofs. It is useful to note here that Appendix H.J summarizes all model-related notation.

H.A Economic Structure

School principals are the decision-makers in our model; their alternatives are to call a male parent first (m), call a female parent first (f), or call neither parent (n). We index decision-makers by $i = 1, \ldots, N$. We take the experiment for a given decision-maker to end when they choose an alternative $j \in \{m, f, n\}$. We assign a decision of n to decision-makers who do not make a call by our exogenously-determined experiment end date. The observables in our experiment are then (1) the choice $y_i \in \{m, f, n\}$ for each decision-maker, (2) the characteristics of the alternative that is shown to each decision-maker, and (3) which parent makes the request.[1]

We assume that decision-makers potentially face different costs, c_i, of making a phone call and this cost does not depend on which parent is called. For instance, some may have inferior technology or be busier than others. We also assume that decision-makers potentially perceive different benefits and costs from choosing different alternatives, and that these are made up of three components: the decision-maker's belief about the value of a response from each parent, the decision-maker's value from calling the parent who initially made contact, and the deterrents they face to calling that alternative.[2] We let $r_{ij}q_{ij}$ denote decision-maker i's subjective valuation of a response from alternative j, where r_{ij} is the belief about responsiveness and q_{ij} is the belief about j's desire for equal decision-making within the household. We let s_{ij} be the value the decision-maker derives from calling the person who reached out to them.[3] Finally, we denote by δ_{ij} any other deterrents to calling alternative j. We assume that each decision-maker i knows c_i, s_{ij} and δ_{ij}, has beliefs over r_{ij} and q_{ij}, and is risk neutral.[4]

H.B Random Utility Model

We construct a random utility model (McFadden, 1974) of decision-maker behavior in which a decision-maker's utility is the difference between the benefits and costs of calling alternative j. For the expected utility maximizer i, the expected utility of calling alternative j is defined as

$$U_{ij} = \mathbb{E}(r_{ij}q_{ij}) + s_{ij} - \delta_{ij} - c_i, \tag{1}$$

where δ_{ij} is positive if factors other than availability beliefs deter decision-maker i from calling alternative j on average. We think of δ_{ij} as a generalization of a distaste parameter, which includes distaste but also other factors not related to beliefs about availability or desire for equal decision-making, such as social norms. This is our basic random utility formulation.

Because calling no one incurs no cost and provides no benefit, we take the utility of calling

[1] In Appendix H.G, we extend the model to incorporate the characteristics of the decision-makers.

[2] We frame this as a deterrent term to align with the distaste parameter in much of the literature. Note that if the decision-maker perceives a benefit from calling a particular alternative, then the deterrent term will be negative.

[3] We include this parameter for two reasons. First, decision-makers in our survey indicated that there is a strong norm around responding to the parent who initiates contact and/or who is listed first on a child's information form; this suggests that violating the norm would negatively impact utility. Second, this was borne out in the data: we saw, treatment by treatment, principals were more likely to call the father when the email came from the father and more likely to call the mother when the email came from the mother.

[4] In Appendix H.I, we discuss relaxing the assumption of risk neutrality. Note, in a previous version of the paper, we presented a slightly different version of the model, which we discuss in footnote 10.

neither parent to be zero. This normalization will play an important role in identification because choice in this context is determined by differences in utility, not levels.

Under this normalization and in our context of choice between calling either of two parents or calling neither parent, decision-maker i calls neither parent if both $U_{im} < 0$ and $U_{if} < 0$; calls the female parent if $U_{if} \geq 0$ and $U_{im} \leq U_{if}$; and calls the male parent if $U_{im} \geq 0$ and $U_{if} < U_{im}$.[5]

We can think of a decision-maker's choice between the three alternatives as having two parts: whether to make a call and which parent to call if they are going to make a call. The cost, c_i, does not affect the decision of which parent to call because the decision-maker incurs the same cost regardless of which parent they call. The cost plays a central role in deciding whether to make a call. In contrast, the choice of which parent to call depends only on the differences in beliefs, the value of replying to the person who sends the email, and other deterrents. To cleanly identify the parameters of interest, we need to consider both the decision of whether to make a call and which parent to call, so we need to include the c_i parameter even if it is not of direct interest.

H.C Experimental Manipulation of Beliefs

Consider an experimental manipulation that sends informative signals to decision-maker i about the availability and desire for equal decision-making of either the female parent ($j = f$) or the male parent ($j = m$). For simplicity, we assume all priors and signals are normally distributed. That is,

$$\bar{r}_j \sim \mathcal{N}(r_j, \omega_j^2), \quad \bar{q}_j \sim \mathcal{N}(q_j, v_j^2), \quad x_{ij} \sim \mathcal{N}(r_j, \sigma_j^2), \quad j \in \{f, m\},$$

where $\bar{r}_j$, $\bar{q}_j$, ω_j^2, and v_j^2 are the prior means and variances common to all i. x_{ij} are signals of the *true* responsiveness r_j of j that we send to i, and the signal variances are σ_j^2.

We assume that the priors for r_f and r_m are independent of the distributions for the equal decision-making, cost, reply-to-sender, and other deterrent parameters. This implies that signals about the availability of a parent (female or male) do not impact the δ_{ij}, s_{ij}, c_i, or q_{ij}. Our assumption that decision-makers are risk-neutral implies that only the marginal means of this distribution are relevant for the expected utility and, therefore, decisions.

Notice that we allow the distributions of the availability signals about the two parents to

[5] We break ties in favor of calling the female parent, but this has no impact in terms of the theory since utility is continuous.

have different means and variances. We also allow for the possibility that signals about one parent may shift the mean beliefs about both parents. This could happen, for instance, if the decision-maker's beliefs about the parents are correlated or if the decision-maker directly infers information about both parents from a signal about just one parent. The impact of a signal about parent j on the decision-maker's belief about the other parent is captured by a correlation parameter ρ_j.

We next describe how decision-makers i update their beliefs after receiving a parental availability signals. To keep the notation simple, we focus without loss of generality on how the belief about the female parent is updated, and the case where the prior belief $\overline{q}_j$ equals one.[6] We then have decision-maker i's posterior means for the responsiveness of parent j as

$$\widetilde{rq}_{if}^{F} = \lambda_f^F \overline{r}_f + (1-\lambda_f^F) x_{if}, \quad \lambda_f^F = \frac{1/\omega_f^2}{1/\omega_f^2 + 1/\sigma_f^2} \tag{2}$$

$$\widetilde{rq}_{if}^{M} = \lambda_f^M \overline{r}_f + (1-\lambda_f^M) \rho_f x_{im}, \quad \lambda_f^M = \frac{1/\omega_m^2}{1/\omega_m^2 + 1/\sigma_f^2}. \tag{3}$$

$\widetilde{rq}_{if}^{F}$ is the updated belief about the female parent after a signal about the female parent, while $\widetilde{rq}_{if}^{M}$ is the updated belief about the female parent after a signal about the male parent. That is, there are two reasons that decision-maker i's belief about the female parent would be updated: directly via a signal about the female parent or indirectly via a signal about the male parent.[7]

Substituting the updated beliefs into Equation 1 gives us the full model equations, which we rearrange into a reduced form that can be estimated directly from our data on which parent sends the email, which availability message is sent, and who the principal calls.

We complete the model by assuming that the errors in each equation are distributed according to the standard Gumbel distribution. This implies that the error differences are distributed according to the standard logistic distribution, helping to simplify the identification argument. Importantly, the random assignment of availability messages to decision-makers implies that the regressors are independent of the errors, so that we can recover the structural parameters—in particular, for prior beliefs, the reply-to-sender motive, and other deterrents—from the reduced-form regression results.

[6] We will discuss signals about desired equal decision making below.

[7] This formulation can be generalized for the case where one sends signals about both parents to the same decision-maker.

H.D Identification

If we only send signals about parents' availability as discussed above,[8] we will be able to cleanly identify $\bar{r}_f$ and $\bar{r}_m$ as well as the reply-to-sender parameters for each treatment, s_j^t for $t \in \{\boldsymbol{noSignal}, \boldsymbol{highFemale}, \boldsymbol{lowFemale}, \boldsymbol{highMale}, \boldsymbol{lowMale}\}$. We will not, however, be able to identify the other deterrent parameters or the updating parameters. The problem is that the effects of beliefs about parents' desire for equal decision-making will be absorbed into these parameters.

To address this concern, we set aside our four signal treatments from the Baseline variation, which contain information only about availability (that is, all treatments except No Signal). Instead, we use the four signal treatments from the Equal Decision variation, adding the statement, "This is the type of decision we both want to be involved in equally" to fix the decision-maker's belief about parents' desire for equality.

If we assume that the value of this signal about parents' desire for equality has a given cardinal value that scales the availability belief and signal, we can cleanly identify the reply-to-sender motive, the joint beliefs $\bar{r}_j\bar{q}_j$ about each parent, the difference between the other deterrents parameters for male versus female parents, the correlation parameters ρ_j, and the weights decision-makers place on their prior beliefs versus the signals, λ_j^J.

Identifying these structural parameters is straightforward, given the four elements of our setting and our model. First, the random utility model provides the structure for the relationship between benefits, costs, and outcomes. Second, calling neither parent provides a clear normalization because it provides no benefits and incurs no costs. Third, experimental randomization establishes that the regressors are not dependent on the outcome variable. Fourth, the assumption that errors are drawn from the logistic distribution leads to closed-form equations for the outcome probabilities.

This would be a standard random utility model if our reduced-form parameters did not vary across the j choices. However, having intercepts and slopes that vary across alternatives is crucial to learning about how experimental manipulation impacts decision-makers choices. Fortunately, the model's structure allows us to identify these intercepts and slopes. Appendix H.D.1 provides intuition for, and proof of, the identification of the reduced-form parameters. We achieve identification by (1) using the proportions of signal-outcome-sender triplets in the data where there are two distinct signals about each alternative $j \in \{f, m\}$ and

[8] We continue to assume that signals about availability do not impact the belief about desired equality so that the prior belief about desired equality is simply carried along with the signal. This is plausible if we conceptually include the ways in which beliefs about desired equality of decision-making impact parental availability in the q_j's.

(2) imposing known cardinal values for each signal. Specifically, we send both positive and negative signals about each parent's availability and assume the values are 1 and -1.[9] We assume the value of the signal about the desire for equal decision-making is 1 to match the value for the high availability treatment. Appendix H.D.2 shows that, with these assumptions, the identification of the structural parameters follows directly from the identification of the reduced-form parameters.

H.D.1 Identification of Reduced Form Parameters

We first combine the economic structure in Section H.A with the random utility model in Section H.B and our experimental manipulation in Section H.C to derived the reduced form of our model. A summary of the crucial assumptions of those sections follows.

1. decision-maker i chooses from among three alternatives: $j \in \{n, f, m\}$.
2. decision-maker i holds probabilistic beliefs about the probability that alternative j will respond to a phone call, $r_{ij} \sim \mathcal{N}(r_j, \omega_j^2)$.
3. decision-maker i holds probabilistic beliefs about the probability that alternative j will desire to be equally involved in the decision, $q_{ij} \sim \mathcal{N}(\bar{q}_j, v_j^2)$.
4. Each decision-maker faces a cost c_i for making a call that is the same for alternatives f and m. c is the population mean of c_i.
5. Each decision-maker has a deterrent parameter for calling that varies by alternative.
6. Each decision-maker has a preference for responding to the parent who sends the message. This preference may depend on the message. We define the variable s_{ij} to be equal to 1 when the female parent sends the message and -1 when the male parent sends the message and we allow for interactions with each treatment, which we denote of s_{ij}^t.
7. Each decision-maker i knows c_i, s_{ij}^t, and δ_{ij}.
8. Decision-makers are risk neutral.[10]
9. Expected utility for decision-maker i is $\mathbb{E}(U_{ij}) = \mathbb{E}(q_{ij} r_{ij}) + s_{ij} - (\delta_{ij} + c_i)$ for $j \in \{n, f, m\}$ with $\mathbb{E}(U_{in}) = 0$.

[9] Appendix H.H discusses the robustness of the results to changes in the assigned values of the signals and/or their symmetry.

[10] We have assumed that decision-makers are risk neutral with respect to the decision about whether and whom to call. In Appendix H.I, we discuss relaxing this assumption.

10. The experimenters choose signal values x^r_{ij} about availability at random to show each decision-maker and send a signal $x_{ij} \in \{-1, 1\}$ about the availability of at most one alternative to each decision-maker. The decision-makers believe that $x_{ij} \sim \mathcal{N}(r_j, \sigma_j^2)$, $j \in \{f, m\}$, where r_j is the true responsiveness of j.

11. A signal x_{ij} can shift the belief $\tilde{rq}_{ij}$ but does not affect c_i, s_{ij}, or δ_{ij}.

12. The experimenters vary whether a positive signal about parents' desire for equal decision making is also sent to a decision-maker. The cardinal value of this signal is the same as the positive signal about availability, that is, 1.

13. ε_{ij} are each distributed according to the standard Gumbel distribution.

We must take a stand on how decision-makers will interpret our signals about availability given that their beliefs also contain the desire-for-equal-decision making component. If the signals about availability and the desire for equal involvement do not interact, the beliefs in Expressions 2 and 3 about females become

$$\widetilde{qr}^F_{if} = \lambda^F_f \bar{q}_f \bar{r}_f + (1 - \lambda^F_f) x_{if}, \quad \lambda^F_f = \frac{1/\omega_f^2}{1/\omega_f^2 + 1/\sigma_f^2} \tag{4}$$

$$\widetilde{qr}^M_{if} = \lambda^M_f \bar{q}_f \bar{r}_f + (1 - \lambda^M_f) \rho_f x_{im}, \quad \lambda^M_f = \frac{1/\omega_m^2}{1/\omega_m^2 + 1/\sigma_m^2} \tag{5}$$

where the superscripts F and M denote the parent about whom the message was sent.

We let w_{ij} be an indicator for sending i a verbal signal of availability (as opposed to the message with no verbal signal) and we substitute these updated beliefs into Equation 1 to get the expected utility from calling the female parent after updating on the signal.

$$\mathbb{E}(U_{if}) = (1 - w_{if} - w_{im}) \bar{q}_f \bar{r}_f + w_{if} \widetilde{qr}^F_{if}(x_{if}) + w_{im} \widetilde{qr}^M_{if}(x_{im}) + s^t_{if} - (\delta_{if} + c_i) \tag{6}$$

$$= (1 - w_{if} - w_{im}) \bar{q}_f \bar{r}_f + w_{if} \left[\lambda^F_f \bar{q}_f \bar{r}_f + (1 - \lambda^F_f) x_{if} \right] + \tag{7}$$

$$w_{im} \left[\lambda^M_f \bar{q}_f \bar{r}_f + (1 - \lambda^M_f) \rho_f x_{im} \right] + s^t_{if} - (\delta_{if} + c_i) \tag{8}$$

$$= \bar{q}_f \bar{r}_f - (1 - \lambda^F_f) \bar{q}_f \bar{r}_f w_{if} - (1 - \lambda^M_f) \bar{q}_f \bar{r}_f w_{im} + \tag{9}$$

$$(1 - \lambda^F_f) w_{if} x_{if} + (1 - \lambda^M_f) \rho_f w_{im} x_{im} + s^t_{if} - (\delta_{if} + c_i) \tag{10}$$

$$= \alpha_f + \eta^F_f w_{if} + \eta^M_f w_{im} + \gamma^F_f w_{if} x_{if} + \gamma^M_f w_{im} x_{im} + s^t_{if} + \varepsilon_{if} \tag{11}$$

where the last equation follows from the mapping below:

$$\alpha_f = \bar{q}_f \bar{r}_f - \bar{\delta}_f - c \tag{12}$$

$$\eta_f^F = -(1 - \lambda_f^F)\bar{q}_f \bar{r}_f \tag{13}$$

$$\eta_f^M = -(1 - \lambda_f^M)\bar{q}_f \bar{r}_f \tag{14}$$

$$\gamma_f^F = (1 - \lambda_f^F) \tag{15}$$

$$\gamma_f^M = (1 - \lambda_f^M)\rho_f \tag{16}$$

$$\varepsilon_{if} = (c - c_i) + (\bar{\delta}_f - \delta_{if}). \tag{17}$$

The ε_{if} are econometric errors and are mean zero because the average terms $\bar{\delta}_f$ and c are absorbed in the constant α_f. Importantly, the random assignment of x_{if} and w_{if} imply that they are independent of ε_{if}. Analogous expressions hold for calling a male parent. Recall that the utility of calling neither parent (U_{in}) is assumed to be zero.

We assume that the ε_{ij} are each distributed according to the standard Gumbel distribution, which implies that the error differences are distributed according to the standard logistic distribution. We next make the identification argument in terms of these econometric errors.

We identify the reduced-form parameters using ratios of the proportions of signal-sender-outcome triplets (which signal was sent, which parent sent it, and which parent—including neither—is called). We denote the proportions as $p_j^{t,J}$. The subscript indicates which parent was called. The first superscript indicates treatments $t \in \{n\boldsymbol{S}, l\boldsymbol{F}, h\boldsymbol{F}, l\boldsymbol{M}, h\boldsymbol{M}\}$, where $n\boldsymbol{S}$ is the No Signal treatment, treatment $l\boldsymbol{F}$ sends the low signal about the female parent, treatment $h\boldsymbol{F}$ sends the high signal about the female parent, treatment $l\boldsymbol{M}$ sends the low signal about the male parent, and treatment $h\boldsymbol{M}$ sends the high signal about the male parent. The second superscript indicates which parent sent the message. For example, $p_n^{l\boldsymbol{F},M}$ is the proportion of decision-makers who receive the low signal about female parent availability from the male parent and then call neither parent.

Given the assumption that $\alpha_n = 0$, the other α_j intercepts are directly identified by comparing the proportions of decision-makers who receive no signal and call parent j and the proportions who receive no signal and call neither parent. To separately identify γ_j^J and η_j^J, we need to create variation in the term $w_{ij}x_{ij}$, that is, the interaction of the indicator variable for whether a signal was sent (w_{ij}) and the value of the signal (x_{ij}). This variation must be distinct from the variation in w_{ij} alone. We achieve this by sending two values of the signal about each alternative j with known cardinal values. Specifically, we send both a positive signal and a negative signal about each parent and assume the values are 1 and -1.[11]

Given the assumptions above and using the no-signal message plus the four availability signal treatments that include the positive signal about parents' desire for equal decision making, we can use the observable proportions of decision-makers for each message-outcome-signal triplet to identify the reduced-form parameters.

Lemma 1. *Given the assumptions of Sections H.A–H.D, the reduced-form parameters α_j, γ_j^J, and η_j^J are identified for $j, J \in \{f, m\}$.*

Proof: We begin with the case in which no signal is sent about either alternative, i.e. $w_{ij} = 0\ \forall j$. Here, the terms involving η_j^J and γ_j^J are zero for all decision-makers, so we have $U_{ij} = \alpha_j\ \forall j$. Because $U_{in} = \alpha_n = 0$ by assumption, the probabilities from the logit model are

$$p_n^{n\boldsymbol{S},F} \equiv \frac{1}{Z^{n\boldsymbol{S}}} \quad p_f^{n\boldsymbol{S},F} \equiv \frac{e^{\alpha_f + s_f^{n\boldsymbol{S}}}}{Z^{n\boldsymbol{S}}} \quad p_m^{n\boldsymbol{S},F} \equiv \frac{e^{\alpha_m + s_m^{n\boldsymbol{S}}}}{Z^{n\boldsymbol{S}}}$$

$$p_n^{n\boldsymbol{S},M} \equiv \frac{1}{Z^{n\boldsymbol{S}}} \quad p_f^{n\boldsymbol{S},M} \equiv \frac{e^{\alpha_f - s_f^{n\boldsymbol{S}}}}{Z^{n\boldsymbol{S}}} \quad p_m^{n\boldsymbol{S},M} \equiv \frac{e^{\alpha_m - s_m^{n\boldsymbol{S}}}}{Z^{n\boldsymbol{S}}}$$

where $Z^{n\boldsymbol{S}} = 1 + e^{\alpha_f + s_f^{n\boldsymbol{S}}} + e^{\alpha_m + s_m^{n\boldsymbol{S}}} + 1 + e^{\alpha_f - s_f^{n\boldsymbol{S}}} + e^{\alpha_m - s_m^{n\boldsymbol{S}}}$. Subscripts denote which alternative is chosen, the first superscript $n\boldsymbol{S}$ denotes that no signal is sent about either alternative, and the second superscript denotes which parent sent the message.

Sending a signal ($w_{if} = 1$) with value $x_{if} = 1$ about alternative f and no signal about alternative m makes the deterministic part of utility for alternative f (i.e. Equation 11 without the error) $\alpha_f + \eta_f^F + \gamma_f^F \pm s_f^{h\boldsymbol{F}}$. We therefore have the following probabilities:

$$p_n^{h\boldsymbol{F},F} \equiv \frac{1}{Z^{h\boldsymbol{F}}} \quad p_f^{h\boldsymbol{F},F} \equiv \frac{e^{\alpha_f + \eta_f^F + \gamma_f^F + s_f^{h\boldsymbol{F}}}}{Z^{h\boldsymbol{F}}} \quad p_m^{h\boldsymbol{F},F} \equiv \frac{e^{\alpha_m + s_m^{h\boldsymbol{F}}}}{Z^{h\boldsymbol{F}}}$$

$$p_n^{h\boldsymbol{F},M} \equiv \frac{1}{Z^{h\boldsymbol{F}}} \quad p_f^{h\boldsymbol{F},M} \equiv \frac{e^{\alpha_f + \eta_f^F + \gamma_f^F - s_f^{h\boldsymbol{F}}}}{Z^{h\boldsymbol{F}}} \quad p_m^{h\boldsymbol{F},M} \equiv \frac{e^{\alpha_m - s_m^{h\boldsymbol{F}}}}{Z^{h\boldsymbol{F}}}$$

where $Z^{h\boldsymbol{F}} = 1 + e^{\alpha_f + \eta_f^F + \gamma_f^F + s_f^{h\boldsymbol{F}}} + e^{\alpha_m + s_m^{h\boldsymbol{F}}} + 1 + e^{\alpha_f + \eta_f^F + \gamma_f^F - s_f^{h\boldsymbol{F}}} + e^{\alpha_m - s_m^{h\boldsymbol{F}}}$ and the superscript $h\boldsymbol{F}$ denotes that we send only a high signal (i.e. value of 1) about alternative f.

Similarly, sending a signal with value $x_{if} = -1$ about alternative f and no signal about alternative m makes the deterministic part of utility for alternative f $\alpha_f + \eta_f^F - \gamma_f^F \pm s_f^{l\boldsymbol{F}}$. We

[11]For a discussion of the impact of the chosen scale of signals, see Section H.H.

therefore have the following probabilities:

$$p_n^{l\boldsymbol{F},F} \equiv \frac{1}{Z^{l\boldsymbol{F}}} \quad p_f^{l\boldsymbol{F},F} \equiv \frac{e^{\alpha_f+\eta_f^F-\gamma_f^F+s_f^{l\boldsymbol{F}}}}{Z^{l\boldsymbol{F}}} \quad p_m^{l\boldsymbol{F},F} \equiv \frac{e^{\alpha_m+s_m^{l\boldsymbol{F}}}}{Z^{l\boldsymbol{F}}}$$

$$p_n^{l\boldsymbol{F},M} \equiv \frac{1}{Z^{l\boldsymbol{F}}} \quad p_f^{l\boldsymbol{F},M} \equiv \frac{e^{\alpha_f+\eta_f^F-\gamma_f^F-s_f^{l\boldsymbol{F}}}}{Z^{l\boldsymbol{F}}} \quad p_m^{l\boldsymbol{F},M} \equiv \frac{e^{\alpha_m-s_m^{l\boldsymbol{F}}}}{Z^{l\boldsymbol{F}}}$$

where $Z^{l\boldsymbol{F}} = 1 + e^{\alpha_f+\eta_f^F-\gamma_f^F+s_f^{l\boldsymbol{F}}} + e^{\alpha_m+s_m^{l\boldsymbol{F}}} + 1 + e^{\alpha_f+\eta_f^F-\gamma_f^F-s_f^{l\boldsymbol{F}}} + e^{\alpha_m-s_m^{l\boldsymbol{F}}}$ and the superscript $l\boldsymbol{F}$ denotes that we send only a low signal (i.e. value of -1) about alternative f.

We repeat each of the last two conditions for alternative m. Sending a signal ($w_{im} = 1$) with value $x_{im} = 1$ about alternative m and no signal about alternative f leads to the following probabilities:

$$p_n^{h\boldsymbol{M},F} \equiv \frac{1}{Z^{h\boldsymbol{M}}} \quad p_f^{h\boldsymbol{M},F} \equiv \frac{e^{\alpha_f+s_f^{h\boldsymbol{M}}}}{Z^{h\boldsymbol{M}}} \quad p_m^{h\boldsymbol{M},F} \equiv \frac{e^{\alpha_m+\eta_m^F+\gamma_m^F+s_m^{h\boldsymbol{M}}}}{Z^{h\boldsymbol{M}}}$$

$$p_n^{h\boldsymbol{M},M} \equiv \frac{1}{Z^{h\boldsymbol{M}}} \quad p_f^{h\boldsymbol{M},M} \equiv \frac{e^{\alpha_f-s_f^{h\boldsymbol{M}}}}{Z^{h\boldsymbol{M}}} \quad p_m^{h\boldsymbol{M},M} \equiv \frac{e^{\alpha_m+\eta_m^M+\gamma_m^M-s_m^{h\boldsymbol{M}}}}{Z^{h\boldsymbol{M}}}$$

where $Z^{h\boldsymbol{M}} = 1 + e^{\alpha_f+s_f^{h\boldsymbol{M}}} + e^{\alpha_m+\eta_m^F+\gamma_m^F+s_m^{h\boldsymbol{M}}} + 1 + e^{\alpha_f-s_f^{h\boldsymbol{M}}} + e^{\alpha_m+\eta_m^M+\gamma_m^M-s_m^{h\boldsymbol{M}}}$ and the superscript $h\boldsymbol{M}$ denotes that we send only a high signal (i.e. value of 1) about alternative m.

Sending a signal with value $x_{im} = -1$ about alternative m and no signal about alternative f leads to the following probabilities:

$$p_n^{l\boldsymbol{M},F} \equiv \frac{1}{Z^{l\boldsymbol{M}}} \quad p_f^{l\boldsymbol{M},F} \equiv \frac{e^{\alpha_f+s_f^{h\boldsymbol{M}}}}{Z^{l\boldsymbol{M}}} \quad p_m^{l\boldsymbol{M},F} \equiv \frac{e^{\alpha_m+\eta_m^F-\gamma_m^F+s_m^{h\boldsymbol{M}}}}{Z^{l\boldsymbol{M}}}$$

$$p_n^{l\boldsymbol{M},M} \equiv \frac{1}{Z^{l\boldsymbol{M}}} \quad p_f^{l\boldsymbol{M},M} \equiv \frac{e^{\alpha_f-s_f^{h\boldsymbol{M}}}}{Z^{l\boldsymbol{M}}} \quad p_m^{l\boldsymbol{M},M} \equiv \frac{e^{\alpha_m+\eta_m^M-\gamma_m^M-s_m^{h\boldsymbol{M}}}}{Z^{l\boldsymbol{M}}}$$

where $Z^{l\boldsymbol{M}} = 1 + e^{\alpha_f+s_f^{l\boldsymbol{M}}} + e^{\alpha_m+\eta_m^F-\gamma_m^F+s_m^{l\boldsymbol{M}}} + 1 + e^{\alpha_f-s_f^{l\boldsymbol{M}}} + e^{\alpha_m+\eta_m^M-\gamma_m^M-s_m^{l\boldsymbol{M}}}$ and the superscript $l\boldsymbol{M}$ denotes that we send only a low signal (i.e. value of -1) about alternative m.

Next, we manipulate the logit probabilities to identify reduced-form parameters $\alpha_j, \eta_j^J, \gamma_j^J$ and s_j^t, which are both reduced-form and structural parameters. As above, we focus without loss of generality on the parameters for calling the female parent.

In order to identify α_f, we take ratios of the probabilities for when no signal is sent.

$$\frac{p_f^{n\boldsymbol{S},F}}{p_n^{n\boldsymbol{S},F}} = e^{\alpha_f + s_f^{n\boldsymbol{S}}} \Leftrightarrow \alpha_f + s_f^{n\boldsymbol{S}} = \ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} \tag{18}$$

$$\frac{p_f^{n\boldsymbol{S},M}}{p_n^{n\boldsymbol{S},M}} = e^{\alpha_f - s_f^{n\boldsymbol{S}}} \Leftrightarrow \alpha_f - s_f^{n\boldsymbol{S}} = \ln p_f^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M} \tag{19}$$

Adding Equation 18 from Equation 19 and then simplifying, we have

$$\boxed{\alpha_f = \tfrac{1}{2}\left[\ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} + \ln p_f^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M}\right]} \tag{20}$$

If we instead subtract Equation 19 from Equation 18 and then simplify, we have

$$\boxed{s_f^{n\boldsymbol{S}} = \tfrac{1}{2}\left[\ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} - \ln p_f^{n\boldsymbol{S},M} + \ln p_n^{n\boldsymbol{S},M}\right]} \tag{21}$$

To identify γ_f^F, we first need to identify $s_f^{h\boldsymbol{F}}$ and $s_f^{l\boldsymbol{F}}$. To do so, we need the following four relationships:

$$\frac{p_f^{h\boldsymbol{F},F}}{p_n^{h\boldsymbol{F},F}} = e^{\alpha_f + \eta_f^F + \gamma_f^F + s_f^{h\boldsymbol{F}}} \Leftrightarrow \alpha_f + \eta_f^F + \gamma_f^F + s_f^{h\boldsymbol{F}} = \ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} \tag{22}$$

$$\frac{p_f^{h\boldsymbol{F},M}}{p_n^{h\boldsymbol{F},M}} = e^{\alpha_f + \eta_f^F + \gamma_f^F - s_f^{h\boldsymbol{F}}} \Leftrightarrow \alpha_f + \eta_f^F + \gamma_f^F - s_f^{h\boldsymbol{F}} = \ln p_f^{h\boldsymbol{F},M} - \ln p_n^{h\boldsymbol{F},M} \tag{23}$$

$$\frac{p_f^{l\boldsymbol{F},F}}{p_n^{l\boldsymbol{F},F}} = e^{\alpha_f + \eta_f^F - \gamma_f^F + s_f^{l\boldsymbol{F}}} \Leftrightarrow \alpha_f + \eta_f^F - \gamma_f^F + s_f^{l\boldsymbol{F}} = \ln p_f^{l\boldsymbol{F},F} - \ln p_n^{l\boldsymbol{F},F} \tag{24}$$

$$\frac{p_f^{l\boldsymbol{F},M}}{p_n^{l\boldsymbol{F},M}} = e^{\alpha_f + \eta_f^F - \gamma_f^F - s_f^{l\boldsymbol{F}}} \Leftrightarrow \alpha_f + \eta_f^F - \gamma_f^F - s_f^{l\boldsymbol{F}} = \ln p_f^{l\boldsymbol{F},M} - \ln p_n^{l\boldsymbol{F},M} \tag{25}$$

Subtracting Equation 23 from Equation 22 and then simplifying, we have

$$\boxed{s_f^{h\boldsymbol{F}} = \tfrac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{h\boldsymbol{F},M} + \ln p_n^{h\boldsymbol{F},M}\right]} \tag{26}$$

Likewise, subtracting Equation 25 from Equation 24 and then simplifying, we have

$$\boxed{s_f^{l\boldsymbol{F}} = \tfrac{1}{2}\left[\ln p_f^{l\boldsymbol{F},F} - \ln p_n^{l\boldsymbol{F},F} - \ln p_f^{l\boldsymbol{F},M} + \ln p_n^{l\boldsymbol{F},M}\right]} \tag{27}$$

Now, if we subtract Equation (24) from Equation (22), we have

$$2\gamma_f^F + s_f^{h\boldsymbol{F}} - s_f^{l\boldsymbol{F}} = \ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{l\boldsymbol{F},F} + \ln p_n^{l\boldsymbol{F},F}$$

Substituting in for the reply to sender terms from (26) and (27), we have

$$\begin{aligned}\gamma_f^F = s_f^{l\boldsymbol{F}} - s_f^{h\boldsymbol{F}} + \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{l\boldsymbol{F},F} + \ln p_n^{l\boldsymbol{F},F}\right] \\ = \frac{1}{2}\left[\ln p_f^{l\boldsymbol{F},F} - \ln p_n^{l\boldsymbol{F},F} - \ln p_f^{l\boldsymbol{F},M} + \ln p_n^{l\boldsymbol{F},M}\right] \\ - \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{h\boldsymbol{F},M} + \ln p_n^{h\boldsymbol{F},M}\right] \\ + \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{l\boldsymbol{F},F} + \ln p_n^{l\boldsymbol{F},F}\right]\end{aligned}$$

Simplifying, we have

$$\boxed{\gamma_f^F = \tfrac{1}{2}\left[\ln p_f^{h\boldsymbol{F},M} - \ln p_n^{h\boldsymbol{F},M} - \ln p_f^{l\boldsymbol{F},M} + \ln p_n^{l\boldsymbol{F},M}\right]} \tag{28}$$

Combining Equations (20), (22), (26) and (28), we have

$$\begin{aligned}\frac{1}{2}\left[\ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} + \ln p_f^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M}\right] + \eta_f^F \\ + \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},M} - \ln p_n^{h\boldsymbol{F},M} - \ln p_f^{l\boldsymbol{F},M} + \ln p_n^{l\boldsymbol{F},M}\right] + \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} - \ln p_f^{h\boldsymbol{F},M} + \ln p_n^{h\boldsymbol{F},M}\right] \\ = \ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F}\end{aligned}$$

We get η_f^F by simplifying the above equation and solving for η_f^F.

$$\begin{aligned}\frac{1}{2}\left[\ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} + \ln p_f^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M}\right] + \eta_f^F \\ - \frac{1}{2}\left[+\ln p_f^{l\boldsymbol{F},M} - \ln p_n^{l\boldsymbol{F},M}\right] = \frac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F}\right]\end{aligned}$$

$$\boxed{\eta_f^F = \tfrac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} + \ln p_f^{l\boldsymbol{F},M} - \ln p_n^{l\boldsymbol{F},M} - \ln p_f^{n\boldsymbol{S},F} + \ln p_n^{n\boldsymbol{S},F} - \ln p_f^{n\boldsymbol{S},M} + \ln p_n^{n\boldsymbol{S},M}\right]} \tag{29}$$

Analogous equations focusing on calls to the male parent identify α_M, γ_m^M and η_m^M as the following:

$$\boxed{\alpha_m = \tfrac{1}{2}\left[\ln p_m^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M} + \ln p_m^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F}\right]}$$

$$\boxed{\gamma_m^M = \frac{1}{2}\left[\ln p_m^{h\boldsymbol{M},F} - \ln p_n^{h\boldsymbol{M},F} - \ln p_m^{l\boldsymbol{M},F} + \ln p_n^{l\boldsymbol{M},F}\right]}$$

$$\boxed{\eta_m^M = \frac{1}{2}\left[\ln p_m^{h\boldsymbol{M},M} - \ln p_n^{h\boldsymbol{M},M} + \ln p_m^{l\boldsymbol{M},F} - \ln p_n^{l\boldsymbol{M},F} - \ln p_m^{n\boldsymbol{S},M} + \ln p_n^{n\boldsymbol{S},M} - \ln p_m^{n\boldsymbol{S},F} + \ln p_n^{n\boldsymbol{S},F}\right]}$$

Thus, the six key reduced-form parameters of interest are identified. ■

Similar algebraic combinations of the logit probabilities identify the η_j^{-J} and ρ_j parameters. We omit these because we do not focus on the cross-parent effects in the data analysis.

H.D.2 Identification of Structural Parameters

Recall Equations (12), (13) and (15), which map the key reduced-form parameters for female parents to the key structural parameters for female parents:

$$\alpha_f = \overline{q}_f \overline{r}_f - \overline{\delta}_f - c \tag{12}$$

$$\eta_f^F = -(1-\lambda_f^F)\overline{q}_f \overline{r}_f \tag{13}$$

$$\gamma_f^F = 1 - \lambda_f^F. \tag{15}$$

Lemma 1 shows that these reduced-form parameters are identified by the various call proportions in our experimental data. We next use the model structure combined with the identified reduced-form parameters to establish the identification of the key structural parameters.

Result 1. *Given the assumptions of Sections H.A–H.D and Lemma 1, the structural parameters λ_f^J, λ_m^J, $\overline{q}_f\overline{r}_f$,$\overline{q}_m\overline{r}_m$, and $\bar{\delta}_m - \bar{\delta}_f$ are identified for $J \in \{f, m\}$.*

Proof: γ_f^F directly identifies λ_f^F as $\lambda_f^F = 1 - \gamma_f^F$ in a simple rearrangement of Equation (15). Once we have λ_f^F, we combine it with Equation (13) to get $\bar{q}_f\bar{r}_f = -\frac{\eta_f^F}{\gamma_f^F}$. Finally, from Equation (12), we have $\bar{\delta}_f + c = -\frac{\eta_f^F}{\gamma_f^F} - \alpha_f$. Analogous equations for the male parent give us $\lambda_m^M = 1 - \gamma_m^M$, $\bar{q}_m\bar{r}_m = -\frac{\eta_m^M}{\gamma_m^M}$, and $\bar{\delta}_m + c = -\frac{\eta_m^M}{\gamma_m^M} - \alpha_m$.

We cannot separately identify $\bar{\delta}_f$ and $\bar{\delta}_m$. However, we can subtract the expression for $\bar{\delta}_f + c$ from the equation for $\bar{\delta}_m + c$ to get $\bar{\delta}_m - \bar{\delta}_f = \frac{\eta_f^F}{\gamma_f^F} - \frac{\eta_m^M}{\gamma_m^M} + \alpha_f - \alpha_m$. ■

Recall that λ_j^J is composed of σ_j^2 and ω_j^2, but these can't be separately identified since we do not have experimental variation for either σ_j^2 or ω_j^2.

We can develop intuition about the model by looking at the relationships between the reduced-form and structural parameters. For instance, start with the expression for the difference in other detterents parameters:

$$\overline{\delta}_m - \overline{\delta}_f = \frac{\eta_f}{\gamma_f} - \frac{\eta_m}{\gamma_m} + \alpha_f - \alpha_m$$

Now rearrange and substitute in the beliefs to get

$$\alpha_f - \alpha_m = \overline{\delta}_m - \overline{\delta}_f + \overline{r}_f - \overline{r}_m. \tag{30}$$

We can interpret this as indicating that the magnitude of the gender inequality (if indeed $\alpha_f > \alpha_m$) derives from the excess deterrents decision-makers face for calling male parents plus their excess belief in the availability of female parents.

Careful examination of the proof of Result 1 makes clear that the identification of the key parameters is not disturbed by a correlation in the belief updating process. This is because identification of those parameters only involves the number of calls to parent j versus neither parent after a signal about parent j compared to the No Signal message. Although we do not focus on the cross-parent effects, allowing for correlation between the beliefs allows one to test whether independence is a reasonable assumption. It also allows the size of the correlation and any potential differences in the updating processes after signals about male versus female parents to be quantified.

H.E Testable Hypotheses

In Section 3.1, we show that there is, indeed, gender inequality in external demands for parental involvement. That is, when there is no signal about availability, the proportion of decision-makers who call the female parent is larger than the proportion who call the male parent.

The structural parameters identified in Section H.D allow us to learn about the sources of this inequality. It may be that decision-makers believe that the expected value of a response from a female parent is higher than that of a male parent; we find support for this mechanism if $\overline{r}_f\overline{q}_f > \overline{r}_m\overline{q}_m$. It is also possible that decision-makers face larger deterrents to calling male parents than to calling female parents; we find support for this hypothesis if $\overline{\delta}_m - \overline{\delta}_f > 0$. We examine these questions in the following section.

H.F Mapping Treatment Effects to Reduced-Form and Structural Parameters

If we include the treatment-specific reply-to-sender terms as the covariates in X_i, it is straightforward to map the coefficients from the treatment effects regression in Equation 1 to the reduced-form parameters from Equation 11. Both are displayed in Table A.3, where we use the no-verbal-signal treatment from the Baseline variation and the four signal treatments from the Equal Decision variation.

The reduced-form regression in Column 2 of Table A.3 is the result of running an unordered logit over decision-maker i's choice to call neither parent (n), the female parent (f), or the male parent (m). Taking calling neither parent as the base case, we have the following equation for calling the female parent when the email comes from the female parent:

$$p_f^{t,F}(x) = \frac{e^{\alpha_f + \eta_f^F w_{if} + \eta_f^M w_{im} + \gamma_f^F w_{if} x_{if} + \gamma_f^M w_{im} x_{im} + s_{if}^t}}{1 + \sum_{k \in \{f,m\}} e^{\alpha_k + \eta_k^F w_{if} + \eta_k^M w_{im} + \gamma_k^F w_{if} x_{if} + \gamma_k^M w_{im} x_{im} + s_{if}^t}}.$$

We also have analogous equations for calling the female parent when the male parent sends the email and calling the male parent when either parent sends the email.

Notice that it matters both which parent is called and which parent the message is about. η_f^F captures the impact of a signal about the female parent on the probability of calling the female parent, while η_f^M captures the impact of a signal about the male parent on the probability of calling the female parent.

The mapping from the reduced-form coefficients to the treatment effects coefficients is simple and intuitive. To be concrete, let's look at the impact of signals about the male parent on the probability of calling the female parent. The reduced-form equation separates this effect into the impact of sending any signal and the impact of the signal's value, which we assume to be 1 or -1. The treatment effects equation separates this effect into the impact of the high signal about the male parent and the impact of the low signal about the male parent. Thus we have $\beta_f^{h\boldsymbol{M}} = \eta_f^M + \gamma_f^M$; that is, the treatment effect from the high signal about the male parent is equivalent to adding together the impact of receiving any signal about the male parent and the impact of the signal value being 1. Similarly, $\beta_f^{l\boldsymbol{M}} = \eta_f^M - \gamma_f^M$; that is, the treatment effect from the low signal about the male parent is equivalent to subtracting the impact of the signal value being -1 from the impact of receiving any signal.

The same relationship holds for each combination of parent called and signal sent: signals about the female parent and the probability of calling the female parent, signals about the female parent and the probability of calling the male parent, and signals about the male parent and the probability of calling the male parent. The two regressions simply decompose the effects of the signals about the male parent in different ways.

H.G Model with Decision-maker Characteristics

Until now, we have assumed that all decision-makers are identical in terms of their observable characteristics. We can easily allow for decision-makers to differ in their beliefs and tastes according to any observable characteristic that is discrete in nature. We are especially interested in whether the decision-maker works at a religious school as this may correlate with holding more traditional gender normative views. To be clear, we do not change the signals that we send to principals in any way. This model extension simply allows the parameters driving decisions to be different for different types of decision-makers. In particular, the signals we send can impact the beliefs of different types of decision-makers differently.

We let g index the discrete categories that make up the decision-maker characteristic. Here, we focus on the type of school at which the decision-maker works at so that $G = \{R, N\}$, where decision-makers at religious schools are denoted by R and decision-makers at non-religious schools are denoted by N.

With decision-maker characteristics, Equation 1 becomes

$$\mathbb{E}(U_{ij,g}) = \mathbb{E}(r_{ij,g}q_{ij,g}) + s_{ij,g} - \delta_{ij,g} - c_{i,g}$$

Each type g of the decision-maker makes their decision as in Section H.C. The signals about parental responsiveness are not differentiated by type of principal, but the signals may have differential impact on the beliefs of different types. We extend the assumptions of Section H.C so that beliefs are independent across types of decision-maker, e.g., that all $r_{ij,g} \sim \mathcal{N}(\bar{r}_{j,g}, \omega_j^2)$ are independent across g.

All beliefs can now be updated separately for each type of decision-maker. For example, we have that decision-maker i of type g has the following posterior mean for the value of a response from the female parent when the female parent sends the message:

$$\widetilde{qr}^F_{if,g} = \lambda^F_{f,g}\bar{q}_{f,g}\bar{r}_{f,g} + (1 - \lambda^F_{f,g})x_{if}, \quad \lambda^F_f = \frac{1/\omega_f^2}{1/\omega_f^2 + 1/\sigma_f^2}$$

Equation (11) becomes

$$U_{if,g} = \alpha_{f,g} + \eta^F_{f,g} w_{if} + \eta^M_{f,g} w_{im} + \gamma^F_{f,g} w_{if} x_{if} + \gamma^M_{f,g} w_{im} x_{im} + s^t_{if,g} + \varepsilon_{if,g}$$

Similarly, equations (12)-(17) become

$$\begin{aligned}
\alpha_{f,g} &= \bar{q}_{f,g}\bar{r}_{f,g} - \bar{\delta}_{f,g} - c_{f,g} \\
\eta^F_{f,g} &= -(1-\lambda^F_{f,g})\bar{q}_{f,g}\bar{r}_{f,g} \\
\eta^M_{f,g} &= -(1-\lambda^M_{f,g})\bar{q}_{f,g}\bar{r}_{f,g} \\
\gamma^F_{f,g} &= (1-\lambda^F_{f,g}) \\
\gamma^M_{f,g} &= (1-\lambda^M_{f,g})\rho_{f,g} \\
\varepsilon_{if,g} &= (c_g - c_{i,g}) + (\bar{\delta}_{f,g} - \delta_{if,g}).
\end{aligned}$$

where $\bar{\delta}_{f,g}$ denotes the average value of $\delta_{if,g}$. Analogous equations hold for calling the male parent.

We then have the following identification result:

Result 2. *Given the assumptions of Sections H.A–H.D and this section, the reduced-form parameters* $\alpha_{j,g}$, $\gamma^J_{j,g}$, $\eta^J_{j,g}$ *and the structural parameters* $\lambda^J_{f,g}$, $\lambda^J_{m,g}$, $\bar{q}_{f,g}\bar{r}_{f,g}$, $\bar{q}_{m,g}\bar{r}_{m,g}$, *and* $\bar{\delta}_{m,g} - \bar{\delta}_{f,g}$ *are identified for* $j \in \{f, m\}$, $J \in \{f, m\}$ *and* $g \in G$, *G discrete.*

Proof: Repeatedly apply the proofs for Lemma 1 and Result 1 for each $g \in G$. ■

H.H Signal Values and Scaling

We have so far assumed that decision-makers take the value of any positive signal to be $x_{ij} = 1$ and the value of any negative signal to be $x_{ij} = -1$. If we change the assumed signal values symmetrically (e.g., both change from magnitude 1 to magnitude 2), η_j does not change but γ_j does. The intuition is as follows: we have not changed whether a signal arrives or not, so the impact of receiving any signal (i.e., η_j) does not change. However, although the signal's value is now assumed to be different, the term $(1-\lambda_j)w_{ij}x_{ij}$ in Equation 11 does not vary with our assumption about the value of x_{ij}. Instead, when we change x_{ij}, the value of $\gamma_j = (1-\lambda_j)$ adjusts to compensate since w_{ij} is simply an indicator for whether any signal is sent. Therefore γ_j is scaled in the opposite direction of the signal value. For instance, if the signals go from magnitude 1 to magnitude 2, γ_j is cut in half. The intercepts, α_j, do not change since they are entirely determined by the proportions of calls when there is no signal.

If we change the assumed value of just one of the signals (e.g., to $+2/-1$ or $+1/-2$), the new γ_j falls between the γ_j for the $+1/-1$ and $+2/-2$ cases. η_j also changes, falling when the positive signal is larger and rising when the negative signal is larger. Any of these changes then ripple through to the structural parameters.In short, as long as we are willing to take a stand on the value of the signals, the structural parameters are identified. However, the identified values of the structural parameters depend on the values we posit for the signals.

H.I Risk Aversion

We have assumed that decision-makers are risk neutral with respect to the decision about whether and whom to call. If decision-makers are instead risk averse with respect to this decision, the prior variance will play a role in the outcome. Importantly, risk-averse decision-makers who are less uncertain about female parents have an additional reason to call female parents beyond their average beliefs.

In terms of the identification of our parameters, what we attribute entirely to the mean of the belief distribution would then be a combination of the mean and the variance if decision-makers are risk averse. In this case, the parameter we estimate for the mean belief about female parents could be larger than the actual mean belief. If, instead, decision-makers are more uncertain about female parents, our estimated belief about the female parent will be smaller than the actual mean belief. The implications for the belief about the male parent mirror these relationships.

To develop intuition for the effect of risk aversion, imagine that a decision-maker holds the same beliefs and has the same reply-to-sender and other deterrents parameters for both parents. This decision-maker will call the parent about whom she is less uncertain; that is, she calls the parent for whom her updated belief variance is smaller. Given a signal variance that is common to both parents, the updated belief variance is lower for the parent about whom the prior belief variance is lower.

H.J Notation

We provide a summary of our notation as a reference.

Subscripts and superscripts

- $i \in I$: decision-maker subscript
- $j \in \{n, f, m\}$: subscript for which parent to call first
- $J \in \{F, M\}$: superscript for the parent who is the sender of the email
- $g \in \{R, N\}$: additional subscript for principal characteristic
- $t \in \{\boldsymbol{noS}ignal, \boldsymbol{h}igh\boldsymbol{F}emale, \boldsymbol{l}ow\boldsymbol{F}emale, \boldsymbol{h}igh\boldsymbol{M}ale, \boldsymbol{l}ow\boldsymbol{M}ale\}$: treatment superscript. When it is only relevant that a message was sent about a particular parent (not whether it was low or high), we use $\boldsymbol{M}$ and $\boldsymbol{F}$

Objects of interest

1. Structural parameters: δ, s, r, q, λ
   - e.g. $\delta_{m,R}$ for the deterrents principals of religious schools face to calling male parent
2. Reduced form parameters: α, η, γ
   - e.g. $\gamma_{m,R}^{h\boldsymbol{F}, M}$ for impact of signal of female high availability ($h\boldsymbol{F}$) on probability that principal from religious school (R) calls male parent (m) when email comes from male parent (M)
3. Reduced-form regressors: w and x do not vary with principal characteristics, so we have $w_{im}^{h\boldsymbol{F}} = 0$ and $x_{im}^{h\boldsymbol{F}} = 0$ for the impact on principal valuation of calling the male parent when they receive a high signal about the female parent
4. Proportions of decision-makers: $p_{m,R}^{h\boldsymbol{F}, M}$ is proportion of principals from religious schools who call male when male parent sends email saying female parent has high availability
5. Coefficients in treatment effects regression: $\beta^{l\boldsymbol{M}}$, $\beta^{h\boldsymbol{M}}$, $\beta^{l\boldsymbol{F}}$, $\beta^{h\boldsymbol{F}}$
   - e.g. $\beta_{m}^{l\boldsymbol{F}, R}$ for impact of low signal about female parent on the probability that a religious-school principal will call the male parent

I External Validity

Type of Household The primary goal of our work is to identify gender gaps in households with two parents, one of whom identifies as female and the other as male. We fully acknowledge that gender identity takes more than two values, but we have started this research with the two ends of the gender spectrum (male and female).

About 98% of US persons identify as either male or female (Census, 2021). The plurality of households with children under the age of 18, 84%, live in a home with two parents, with 99% of these being opposite-gender couples (Census, 2022).

We believe the direction of the effects of our high/low-availability messages would be the same for various genders (e.g., two non-binary parents, same-sex couples). However we would expect No Signal inequality to be closer to zero in households with these gender identity sets. Nationally representative data indicates that same-sex households do not report wishing they were contacted more/less than they actually are by their child's school.[12]

School Setting Our experiment takes place in a K–12 school setting which we chose because over 40% of households in the US, have school-aged children (NCES, 2021). Almost all parents, 97%, of parents send their children to school outside the home (NCES, 2021). Additionally, the gender gap in time spent on children in school-related activities closely mirrors the overall tendency for mothers to engage in more child-related tasks than fathers (BLS, 2021).

We believe that any gender gaps that we document in our specific task in the school setting will generalize to other tasks in the school setting, such as picking up a sick child, or joining the Parent Teacher Association. Educators in our survey report that they would favor contacting the mother first in many of these scenarios (we discuss the survey in Section M.A). The gender distribution of these tasks is significantly skewed with mothers comprising almost 90% of Parent Teacher Association members and many surveys finding fathers self-report lower levels of involvement in their child's school activities, compared to mothers.[13]

[12]See https://csed.byu.edu/american-family-survey for evidence from 219 respondents who are in a same-sex married couple that is living together and are from a nationally representative sample. The limited survey evidence we have on non-binary parents from this survey does indicate that the four non-binary respondents report being contacted 77% but wishing to be contacted only 60% of the time.

[13]See our own survey in section M.C and Daly and Groes (2017) https://archive.nytimes.com/parenting.blogs.nytimes.com/2009/01/06/dads-in-the-pta/, https://education.gov.scot/media/b3cn2mv5/nih327-dads-involvement-in-school.pdf

J Ethics

There are pre-existing observational data and survey data that shows decision-makers prefer to call mothers more than fathers. However, in this observational data, it is not possible to tell whether mothers have signalled they would like to be called more often. To measure if there is bias without such signals, we need an experiment like the one we have performed. Additionally, in observational data it is difficult to assess the reason that any bias towards calling mothers exists without exogenous variation in the signals being sent by the household about male versus female availability. For both these reasons, an experiment is needed to cleanly identify mother preference without signals, and how much of that bias is driven by signals about availability.

However, experiments come with costs. A common critique of audit studies, which perform outreach from fictitious persons to a third party (often a hiring business), is that the person who receives the message wastes time and effort in evaluating the message. The median time spent leaving our parents a message was less than one minute, with the 99th percentile being a message of less than two minutes. As such, each principal in our dataset is not spending a large amount of time being in our study. Unlike a resume audit study, the principals in our study do not need to evaluate a lengthy fictitious candidate's resume for a position; rather, they need only to read our brief email message and return our call (only 20% of principals call us, and only 17% leave a voice mail, further reducing the likelihood of significant harm to our subjects).

Another concern might be the number of individuals who were contacted. Using our pilot data as a guide, we simulated possible outcomes of samples of varying sizes and chose the smallest sample size the simulations indicated was needed to identify the deep parameters of the theoretical model. This was 80,000 principals out of a total of more than 100,000 in the database of principals.

As a first step toward compensating schools for their time we have donated a total of about $5,000 to the following school- related non-profits and projects: Kids in Need, First Book, Generation Teach, and 10 projects on DonorsChoose.org.[14]

Also, our subjects are school officials who aim to improve school quality as part of their

[14] This type of compensation is non-standard for audit studies. We tried to inform our choice of the amount as follows. Let us assume the educators who responded to our message spent about 20 seconds on reading and responding to our messages. The median school principal salary is $113,000 per year that is a per minute wage rate of $0.015 per second assuming a 40 hour work week and working 52 weeks of the year. That would be 20 second * ($0.015 per second)=$0.30 per school. We were contacted by a total of 15,881 schools and at $0.30 each that is $4,764.30.

position. Our research, in part, informs ways to improve school quality through better serving parents, and as such, participation in our study is arguably part of our subjects' regular job duties.

Subjects were told two weeks after our initial emails that the household no longer needed to talk, thus releasing the subjects from the need to think about the fictitious household. We decided not to debrief our subjects even though debriefing may have the positive aspects of transparency and the ability to withdraw from the study. Here, we followed the logic outlined in Pager (2007):

> The issue of debriefing subjects following the completion of the audit study is a complicated one. Though typically IRB protocol supports the debriefing of subjects whenever possible, in certain cases acknowledging the occurrence or nature of a research study is deemed undesirable. It could be argued, for example, that subjects could be placed at greater risk should their behavior, as a result of the audit study, fall under greater scrutiny by superiors. For human resource personnel or managers who are thought to be discriminating, the consequences maybe more serious than if no attention were brought to the audit whatsoever. While the chances that negative consequences would result from this research in any case are very small, some IRB committees take the view that eliminating the debriefing stage is the most prudent strategy. The purpose of audit research is not to harm individual employers. Rather, the research seeks to improve our understanding of the barriers to employment facing stigmatized groups in their search for employment.

A second concern is that the decision-makers' involvement may harm other non-fictitious persons because of their involvement in the audit study. For example, if a firm decides to call back a fictitious applicant in an audit study, this may crowd out a call to a real applicant. We do not believe our study poses this harm. The act of calling one family likely does not crowd out calling another family.

An additional possible hazard in a labor-market audit study is that the fictitious applicants never accept the job interviews. If they have some identifiable factor, such as foreign-sounding names, this may cause firms to negatively update their views of real persons with foreign-sounding names. Again, we do not think our study poses this harm as all of our households are two-parent households with racially neutral names, as such it is difficult to identify which subgroup a school principal would negatively update about in our study.

Lastly, a large survey of economists finds that researchers are quite comfortable with the lack of informed consent common in natural field experiments like audit studies (Charness et al., 2022). The same survey finds that economists prioritize avoiding more explicit deception but believe it is acceptable for important questions when alternative research designs are unavailable. Informed consent is ideal, but it is difficult to study gender discrimination with informed consent without possibly biasing the results. Recent studies find that informing people they are in a study leads to lower measures of discrimination (Agan et al., 2023). Our study was approved by the relevant Institutional Review Boards (IRBs) at our home institutions, and as such the harms and benefits have been evaluated by a third party that approved the research design.

K Efficiency

Multiple parties are involved in the interaction that we investigate: the parents, the external decision-maker (in our case, the school), the child, and the parent's employers if employed. With multiple parties involved and many trade-offs to consider, it is not readily apparent what the most efficient allocation of calls between mothers and fathers is.

Parents. The existing skew toward mothers contributes to gender gaps in a wide range of labor market and educational outcomes, including career trajectory, occupational choice, and earnings. Workday disruptions stemming from child-related interruptions have also been linked to declines in women's physical and mental health (Zamarro and Prados, 2021). Furthermore, contacting the person the household indicates has more availability would likely reduce parents' stress levels; such reductions in stress are associated with better parenting (Conger et al., 2010).

In our experimental data, even when the the father sends the email and signals that he is highly available, 12% of the calls are still directed to mothers (Table 2). This indicates that households that want a more egalitarian division of child-related tasks and household labor, specifically fathers who want to be more involved, may be limited in achieving their goals in this area. Therefore, the current inequality in demands for parental involvement appears to be inefficient for some parents.

Finally, even if we assume that men and women, *on average*, have different comparative advantages, there is a distribution of these skills within each gender. This implies that households differ from the population average, resulting in a deadweight loss due to inefficiencies within households. This further suggests that reducing the restrictions placed on households by institutions would decrease the deadweight loss.

External decision-makers. Decision-makers may have multiple competing objectives. In our model (Section H), the decision-maker maximizing the likelihood of a useful response–a short-run outcome. However, in the long run, an entity (school, church, extracurricular program, doctor) may find it desirable to have a more diverse set of parents involved (e.g., not skewed toward mothers). They may also prefer to have more parents (e.g., both parents versus one) involved (Clark et al., 1980). A less myopic decision-maker may want to call the father even if they believe he is less likely to respond or may provide a less useful response. We believe investigating these trade-offs is an important area for future research.

Child. The skew toward mothers being called more may be welfare-harming for children, given the extensive evidence that children benefit from having both fathers and mothers involved (Pleck, 2007; Nakata, 2023). Yet, research on the engagement of fathers in child-related social services has found that along with gendered and cultural factors that support a preference for the mother, the institutional aspects of social services result in partial or full exclusions of fathers from child-related interventions (Perez-Vaisvidovsky et al., 2023). This implies considerable welfare costs for children.

Parents' Employers and Economic Efficiency. Parents' employers would like to minimize interruptions to their employees' workday. If the school is going to contact a parent, each employer would prefer that the school contacts the parent it does not employ. This has the flavor of a zero-sum game between the two employers. However, it would be most efficient, from the standpoint of both the mother's and the father's workplaces (and the overall economy), for the parent who has signaled more availability to be contacted, provided that the household has information about which parent is a more productive worker. This would protect the more productive worker's time, increasing the combined output from the two parents. We find evidence that decision-makers listen to these signals but do not fully integrate them, as 26% of the calls still go to mothers even when the father states he is highly available (Table 1).

An important next step in this research agenda is further investigation of the trade-offs each party faces and how a social planner might weigh the needs of the various parties.

L Data Collection and Matching

Emails and Phone Numbers To record phone metadata and voicemails we used a service called Callfire to set up a series of different phone numbers for our male and female parents. First, we set up a series of phone numbers with a generic voice mail and auto-reply text messages saying that the number did not receive text messages. We also set up email addresses with an auto-reply directing responders to please call instead of emailing. The exact email addresses from which we sent our messages were "erica@miller-family.net" and "roy@miller-family.net" for part of our data collection. We switched to emails from "audrey@the-johnsonfamily.net" and "curtis@the-johnsonfamily.net" for the bulk of data collection. We discuss the choice of exact names in detail below and in Section L.A. Due to constraints on email send limits, the follow-up emails sent after the first email which said the family no longer needed to talk were sent from "audrey@the-johnson-family.net " and "curtis@the-johnson-family.net."

Email is a common way for parents to contact schools. In our own survey, three-fourths of educators reported being contacted by parents via email at least once a month (Section M.A). These educators also reported that, when being emailed by both parents, a single parent emailing and cc'ing the other parent was more common than emails from a joint family email account. In one of our pilot data-collection efforts, we found that emailing from a joint email account lowered callback rates (Section L.A). Furthermore, we were concerned that a joint family email address might signal a more egalitarian family, which might bias our results towards finding more equal calls to mothers and fathers. As such, we decided not to use any joint family email accounts.

Names We chose the names from the top 200 listed by the Social Security Administration in 1980 (SSA, 2022a). We chose 1980 because we primarily contact schools that enroll children ages 5 to 18, the average age being 11.5 years old. A child who is 11.5 in 2021 was born in 2009 (2021-11.5=2009.5). The average age of a first-time parent in 2009 was 29.4 years old (CBS, 2019), which means our parents on average would have been born in 1980 (because 2009-29.4=1979.6). From the 1980 list, we chose first names that did not have a strong indication of a specific race or ethnicity (Tzioumis, 2018) (Erica and Roy) and we chose our last names (Johnson and Miller) from the list of the most common last names in the US over many decades (SSA, 2022b). We also did online searches for the names (Audrey Johnson, Curtis Johnson, Erica Miller, Roy Miller) to see if there were any famous or infamous people with these names that might bias our results. In addition we did a Google image search for these names to ensure they encompassed a balance of race and ethnicities.

Messages We pretested our messages using a survey run on Amazon Mechanical Turk to select which messages gave the widest variation in the self-reported likelihood of getting a callback. We also pretested our messages with a set of educators (see Section M.A) to ensure

Table L.1: Longer Versions of Messages

Variation & Treatment	Body Text
Baseline No Signal (Used in Study)	We are searching for schools for our child. Can you call one of us to discuss?
Baseline No Signal (Longer Alternative)	I'm Curtis[Audrey] Johnson. I'm writing to request information about your school because we are searching for schools for our child, Riley. Riley is a well behaved student, and loves most subjects. We're not totally sure when we will be needing to enroll, but we are looking forward to hearing more from you at your earliest convenience. Could you call one of us to discuss? Thank you very much,
Equal Decision No Signal (Used in Study)	We are searching for schools for our child. Can you call one of us to discuss? This is the type of decision we both want to be involved in equally.
Equal Decision No Signal (Longer Alternative)	We are searching for schools for our child. Could you call one of us to discuss? You can call either me or my wife, Audrey [husband, Curtis]. Since we make these kinds of decisions together, whoever you call will convey the information to the other parent. Thank you very much,

the messages seemed natural to this audience. Furthermore, we tested different versions of the two message variations we sent the most (Baseline and Equal Decision). The messages we sent were brief by design in effort to use less of the decision-maker's time and to make our treatments about parent availability more salient. We did test longer versions of our two most-emailed messages, as detailed in Table L.1, but found that the difference in the callback rates was not statistically significant, nor was the proportion of calls to mothers versus fathers.

L.A Pilot Studies

In May of 2021, we sent 767 emails; in June 2021, we sent out 1,250 emails; and in November 2021, we sent out 1,250 emails. The primary purpose of this early data collection was to refine the process by which we send emails, learn about response rates, and test our ability to match phone calls to emails sent. As such, we concentrated on a subset of our treatments: No Signal, Male High Availability, Male Low Availability in the May and June 2021 waves, and expanded to five treatments in the November 2021 wave with the inclusion of the Female High Availability, Female Low Availability treatments.

Our pilot studies tested several procedural items. For our May pilot, we chose the names Jennifer and Michael because they signal gender well. However, Jennifer and Michael are predominantly white names, so we wanted to test a more race-neutral set of names (Erica/Roy) to see if this impacted callbacks. Testing Jennifer/Michael vs. Erica/Roy, we found that using the more race-neutral names (Erica/Roy) decreased callbacks by 8.8 percentage points. We felt that using the more race-neutral names increased the external validity of our findings and as such decided to use them in our full data-collection effort.

Additionally, we tested two types of email accounts in our pilots, given that our survey of educators indicated that the use of a joint family email address was less common than the use of individual email addresses and cc'ing the other parent (Section M.A). We found that using a joint family email address (versus individual email addresses, with one parent cc'ing the other parent) decreased our callback rates by 9.2 percentage points ($p = 0.032$). With the evidence from both the pilot and the survey, we dropped the joint family email address in our full data-collection efforts.

L.B Phone Call Data

May 2022 Phone Calls In May of 2022, we sent about eight thousand emails to schools. However, we found that some of these schools shared a single email address or a single phone number (e.g., a network of charter schools, or a school district that uses a centralemail address and/or central phone number). In addition, an error in our code meant we mistakenly sent more than one email to some email addresses. Removing all these from our dataset, we retained 7,935 emails sent to schools that each had a unique email and unique phone number.

In the weeks following, we received 2,990 callbacks to our May 2022 emails. Some of these callbacks are problematic: some are assuredly in response to emails we dropped from our dataset for the reasons outlined above, and a small number are likely spam calls made to our fictional parents' numbers (though these are most likely randomly distributed across our phone numbers). More of an issue is that these callbacks include calls made by the same school principal using multiple different phone numbers or just calling the same household multiple times in a row to the mother, the father, or some combination of both. Our outcome variable of interest is the first parent contacted, rather than the total number of calls made by a principal (although this could also be of interest). Furthermore, to be able to perform an analysis of a school or principal's specific demographics, we need to link each phone callback to a specific email sent. This matching is a multi-step process.

July 2022 Phone Calls In July and August of 2022, we sent 72,136 emails. In the weeks following, we received 30,214 calls. Much like our May data, these calls include spam calls. Our primary objectives with matching callbacks to specific schools are to allow analysis by the school's attributes and to correctly identify which parent was called first if calls were made from multiple phone numbers by the same school principal.

Matching Phone Calls To Emails First we created a dataset with a single line for each unique phone number. We also included all the phone calls from “Restricted” phone numbers, as it is impossible to tell if those are unique. In May 2022 the one-call dataset had 1,684 lines, and in June/July 2022 the one-call dataset had 17,139 lines. We then matched these CallFire 10-digit phone numbers to the 10-digit phone numbers associated with our schools. A little over 60% of calls matched up.

We then took the remaining CallFire phone calls and performed a “fuzzy” match on the first 6 digits of each phone number. For example, all calls originating from Tufts University start with these same 6 digits, 617-627; all calls from Brigham Young University start with 801-422. We then had research assistants check these fuzzy matches for accuracy and disambiguation when two-plus schools matched to a single CallFire phone call. Around one-fifth of calls are matched by a “fuzzy” match. For the remaining CallFire phone calls, we asked research assistants to listen to voicemails and perform Web searches to attempt to match them to a school we emailed. Last, we randomly selected a subset of these matches to be audited by a different research assistant to check for the quality of our matching.

M Survey Evidence

We collected data for this project via survey three times in 2022, 2023 and 2025. Here we describe those surveys in more detail. All the surveys were run on Prolific (IRB number STUDY00002608).

M.A Educator Survey

In April 2022, we ran a survey of educators before conducting our main field study. People were eligible to take our survey if they were over 18, resided in the US, and answered ”Yes” to the following question “Do you regularly reach out to parents as part of your job (or for a previous job)?” To limit our sample to educators, on the survey platform Prolific we only allowed participants who listed their employment sector as “Education & Training.” We had 238 educator respondents in 2022.

The goal of this survey was to check that the type of email we were sending to schools was appropriate. Over 50% of educators reported getting the most questions about school enrollment during the month of August. August was followed by the months of May, September, July, June, and April (in that order), with about 18% to 28% of educators stating they got the most questions about enrollment in these months. About three-fourths of educators said that being contacted by parents was either very common (at least once a week) or somewhat common (at least once a month). When being emailed by both parents a single parent emailing and cc'ing the other parent was more common than emails from a joint family email account. Educators reported they contacted parents by phone about the same amount as they did via email, email being slightly more common.

A second goal of our survey was to see how educators self-reported calling mothers versus fathers in response to different types of inquiries. We found that educators self-reported they would make no call in response to a message like our Baseline No Signal only 8% of the time; this is very different than the rate we observe in our natural field experiment which is closer to 80% not calling back either parent. This could be because some of our email messages are going to spam, or because the group of survey respondents is a selected group, or because educators are overly confident in their likelihood of making a call. This disconnect highlights the importance of running a natural field experiment in this setting. Interestingly, conditional on self-reporting making a call the educators said they would call the female parent 57% of the time, which is quite similar to the rate we observe in the natural field experiment (Table 1 Panel A.ii Column 3 and Panel B.ii Column 3).

We found that educators always reported a higher level of wanting to contact the mother instead of the father if they had to choose a single parent to contact about a child being sick (98% contact mom), volunteering at a book fair (96%) or career day (78%), school related payments (86%), or a child's allergies (97%). We allowed the educators to rank the following reasons for choosing to contact the person which were displayed in a random order: I expect this person to be more likely to respond quickly, I expect this person to be more likely to be the primary decision-maker about this topic, I simply like interacting with this person more, and Other. The reasons of "I expect this person to be more likely to respond quickly", "I expect this person to be more likely to be primary decision-maker about this topic" were very similarly ranked as the top choice within each type of inquiry.

M.B Decision-Maker Survey

In April 2023, we ran a similar survey of adults who interact with parents, including educators. People were eligible to take our survey if they were over 18, reside in the US and regularly reach out to parents as part of their job. Specifically, they had to answer in the affirmative this question "Do you, or have you ever, regularly interacted with parents?" We had 377 respondents from a variety of persons who interact with parents (the most common were Teacher, Childcare provider, Medical Practitioner, Nurse, Sports Leader). Of the

377 respondents in 2023, 77 self-identified as interacting with parents in the role of "other", and we also include decision-makers with small numbers in this other category. Specifically Religious Leaders (7), Law Enforcement (6) and Family (14) are included in "Other." The primary purpose of this survey was to produce panel B of Figure 1. We randomly assigned whether a decision-maker was asked the following question about a [mother] or [father]: What proportion of the time do you contact the [father][mother] first if only contacting one parent first?

We also asked some of the questions we had asked in our 2022 educator survey of all types of decision-makers. Trends were broadly similar for educators asked in 2022 and 2023, and for educators versus all types of decision-makers.

In Figures D.1 and D.2 we break out our natural field experiment results by the gender of the principal at the school. We can do a similar exercise with our survey data where we compute the average self-reported calls to mothers within each type of decision-maker, and we find a positive correlation as shown in Table M.1.

Table M.1: **Mean Self-Reported Percent of Calls to Mothers By Decision-Maker Type**

Decision-Maker Type	**Proportion Female**	**Contact Mom 1st**	**N**
Childcare provider	0.77	0.78	44
School Related	0.62	0.75	63
Medical/Dental	0.60	0.72	77
Extracurricular	0.47	0.74	30
Sports Leader	0.38	0.58	32
Tutor	0.30	0.70	27

Last, within our surveys, we also identified which respondents were parents from a household with one male and one female parent. In April 2022, there were around 90 respondents who answered a series of questions about households and schools for us; in April 2023, just over an additional 125 parents answered questions about schools and other points of contact (e.g. Doctors, Law Enforcement, Sports). We asked these respondents a number of questions about their experiences as parents, which informed our next survey of households.

M.C Household Surveys

In February 2025, we ran a survey of individuals over 18 who were based in the U.S. and identified as current or former parents in a two-parent household where both parents were present.[15] We had 353 respondents, 45% of whom identified as mothers, and the rest identified as fathers. One purpose of this survey was to measure how child-related interruptions impacted mothers' and fathers' labor market and human capital decisions. We report our findings in Figure 5. We emphasized respondents should think of "non-routine/unexpected child-related interruptions to your job(s)" by external organizations when their children were living at home with them that were initiated by the external organization (for example, a call about a sick child, an email/text to schedule a doctor/dentist appointment, a reminder to register for camp/practice/extracurricular activities).

We compare our sample to the general US population who are married and living together in households with at least one child under the age of 18. Our respondents are slightly younger (43 vs. 55 from IPUMS), more white (78% vs. 54% in IPUMS), more black (11% vs. 6% in IPUMS), less Hispanic (5% vs. 19% in IPUMS) and less Asian (4% vs. 14% in IPUMS), more liberal (30% vs. 21% in NORC), more conservative (46% vs. 29% in NORC), less moderate (24% vs. 50% in NORC), and more likely to hold a BA as their highest level of education (39% vs. 24% from IPUMS).[16] Also, of our female respondents 27% are currently a stay-at-home mother and of our male respondents 3% are currently a stay-at-home father which is quite close to the statistics reported from Pew Research (26% moms and 7% dads) and IPUMS (26% moms and 8% dads).[17] We also ask if our female and male respondents have ever identified as a stay-at-home parent for more than one year while their children were below 18 and living at home. We find that 68% of mothers and 14% of fathers identified as a stay-at-home parent for more than one year while their children were below 18 and living at home, however we have been unable to find nationally representative data which asks a similiar retrospective question.

We also used this survey to better understand how two-parent heterosexual households perceived their interactions with decision-makers at schools and other organizations. We report the findings from this survey throughout the paper to inform our understanding of how a mother vs. father feels about: outsourcing an interruption to their partner, schools/organizations' ability to honor a family's request about who to contact, how often schools/organizations

[15]We also ran versions of this survey in April 2024 ($N = 349$) and September 2024 ($N = 142$). The results from previous versions of this survey are quite similar to the most recent wave of the survey (available from the authors upon request). We ran the most recent wave to be able to address comments brought up during the editorial process.

[16]https://usa.ipums.org/usa/ and https://gssdataexplorer.norc.org/home

[17]IPUMS does not have data on if someone is a current stay at home mom or stay at home dad. With the numbers I provided, I defined the current stay at home mom as the percentage of females with children 0-18 working 10 or less usual hours worked per week, and the current stay at home dad as the percentage of males with children 0-18 working 10 or less usual hours worked per week. Please let me know if you would like me to adjust this definition. The IPUMS data are from 2022. https://www.pewresearch.org/short-reads/2023/08/03/almost-1-in-5-stay-at-home-parents-in-the-us-are-dads/

contact each parent and how often each parent *wishes* the schools/organizations called them.

Additionally our survey allows us to ask mothers and fathers how much they wish they were contacted by individual entities and overall by all entities. One might think that a household hoping to achieve an approximately equal split of contact over all entities might obtain this by assigning the mother as the sole contact for some entities (e.g. school, medical), and the father as the sole contact for others. For those households who wanted an approximately equal split over all entities, we show the reported desired contact from each entity in Figure M.1.[18]

Even when we do not subset on those who wanted an approximately egalitarian split as in Figure M.2 we still see relatively little mass at the 0 or 100 end of these sub-figures. However, there are some notable exceptions with mostly female respondents reporting they want 100% of the contact from tutoring, medical/dental, extracurricular, school and child care providers.

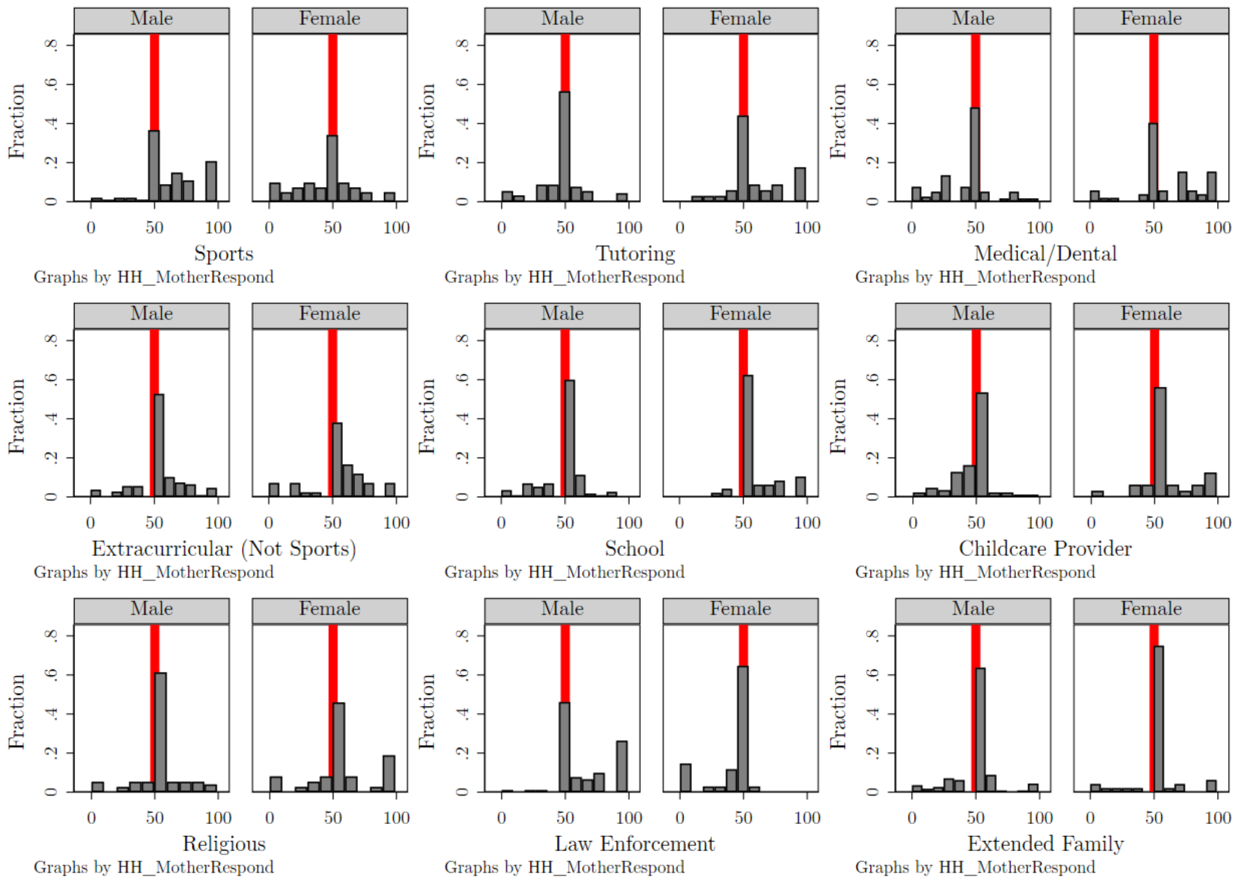
M.C.1 Hypothetical Choice Experiment

To better understand how two-parent households react to the findings from our paper we ran a hypothetical choice experiment without our household survey in February 2025. The experiment was pre-registered at AEARCTR-0015108. We used a between-subjects design where we test the effect of showing the findings of our study on the likelihood a respondent will report that their household would increase the proportion of contact from the father to the school on a five point scale (Decrease A Lot / Decrease Somewhat / No Change / Increase Somewhat / Increase A Lot). Specifically we ask: How does all the information above change the proportion of contact by the father in your household (versus the mother) to the school? Table M.2 shows the messages shown in our main control and treatment arms.[19]

[18]We define those who wanted an approximately equal split over all entities as those who answered 50-50, 40-60 or 60-40 to a question asking: *Consider the time spent being contacted by all the organizations combined (for example, school, child care center, doctor's office, sports coaches, extracurricular activities). What proportion of that time would you and your partner want to be contacted by all those organizations combined?*. This was 49% of our sample.

[19]We also had a version of the control which was slightly shorter than our main treatment message, but had less extraneous information. It read: *A recent study finds that when parents become more involved in their children's school lives by receiving regular information about their child's academic progress, their children develop more positive behavior in school.* In addition we had an arm of the study which showed a message about the effects of our mother low availability message and read: *A recent study finds that when parents become more involved in their children's school lives by receiving regular information about their child's academic progress, their children develop more positive behavior in school. Another recent study finds that if the child's father emails the school and indicates that the mother has limited availability, about 90% of responses from the school will be directed to the father.* We find our main outcome variable of increased contact from the father is not statistically significantly different across our two control messages ($p = 0.136$) nor across our two treatment messages ($p = 0.400$). We pool the control and

Figure M.1: Histograms of Self-Reported Contact Wanted By Mothers and Fathers for Households Which Want Egalitarian Split Over All Entities



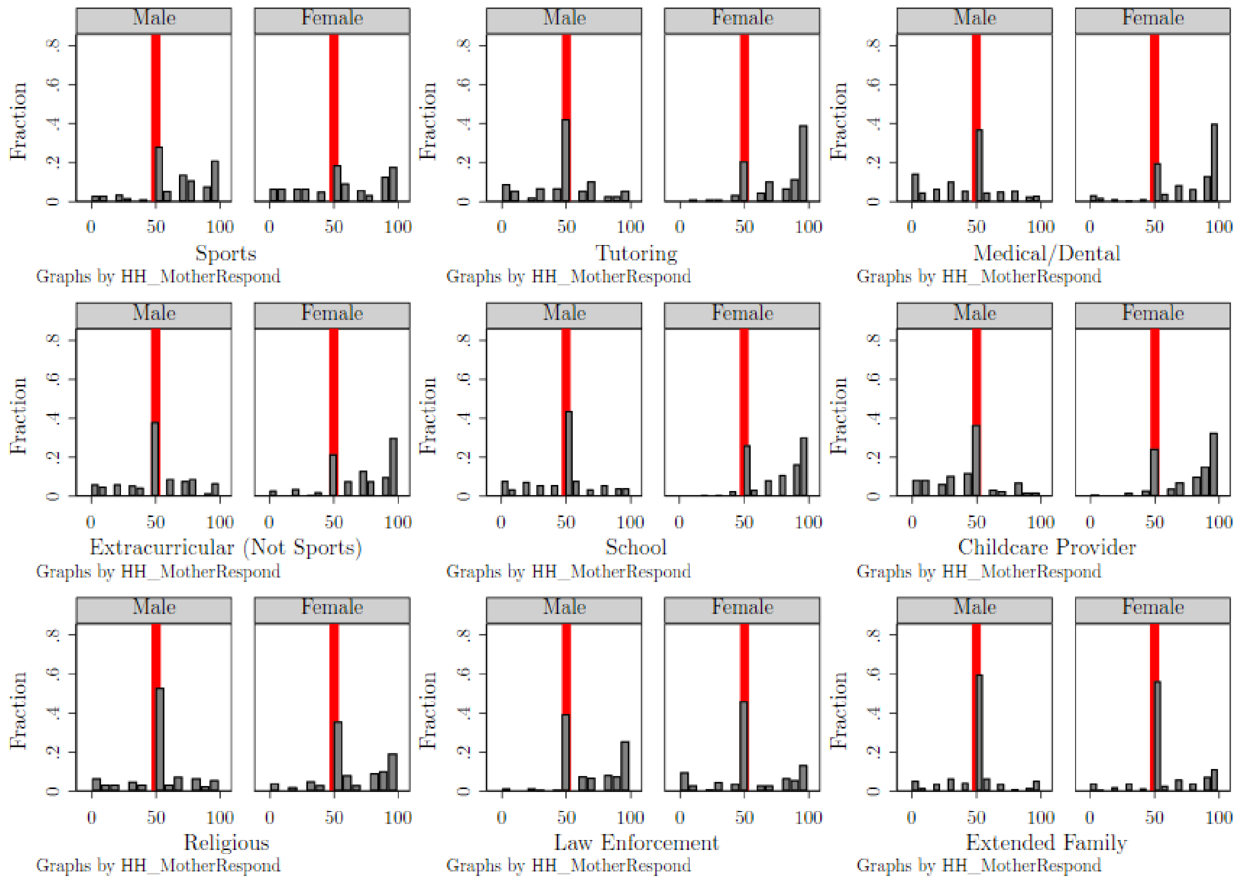
Notes: In this Figure we show histograms of the responses to the question: *We are going to list some common organizations below. How much do you wish each of these organizations contacted you vs. your partner? If your household does not interact with a specific type of organization please select "non-applicable".* There is a vertical red line where for those who answered they wanted an exactly 50-50 split for an entity. We show the answers only for the subset of "Egalitarian" households, meaning those who answered 40-60, 50-50, or 60-40 to the following question: *Consider the time spent being contacted by all the organizations combined (for example, school, child care center, doctor's office, sports coaches, extracurricular activities). What proportion of that time would you and your partner want to be contacted by all those organizations combined?*. This represents 49% of respondents (173 of our 353 survey respondents). Figures exclude those who marked NA for an entity with the following total observations per sub-graph: Sports = 142, Tutoring=126, Medical/Dental=170, Extracurricular (Not Sports) = 150, School=163, Childcare Provider = 118, Religious = 112, Law Enforcement = 125, Extended Family = 161.

We create a binary variable which takes the value one if a respondent states their household will Increase Somewhat or Increase A Lot as our main outcome from the experiment.

After we present these experimental messages, we ask respondents whether they knew the information from the studies beforehand. We found that around 20% of respondents knew the information from the our study in both Treatment and Control ($p = 0.2965$ Control vs. Treatment). The main thing we take away from this is that there may be a lack of information

treatment variations with the main versions in our analysis.

Figure M.2: Histograms of Self-Reported Contact Wanted By Mothers and Fathers for All Households Split Over All Entities



Notes: In this Figure we show histograms of the responses to the question: *We are going to list some common organizations below. How much do you wish each of these organizations contacted you vs. your partner? If your household does not interact with a specific type of organization please select "non-applicable".* There is a vertical red line where for those who answered they wanted an exactly 50-50 split for an entity. We show the answers all of our 353 survey respondents. Figures exclude those who marked NA for an entity with the following total observations per sub-graph: Sports = 284, Tutoring=231, Medical/Dental=347, Extracurricular (Not Sports) = 308, School=335, Childcare Provider = 240, Religious = 219, Law Enforcement = 248, Extended Family = 333.

about our findings insofar as about 80% of respondents do not say they knew about the findings beforehand. We can also use this as a control variable in our analysis of the effect of our Treatment versus Control messages.

Specifically, we find that 41% say they will increase the proportion of contact from the father when told about our study while this is only 30% for those seeing the Control message ($p = 0.0227$). If we limit to only those for whom our study findings provided new information we find the gap widens with only 26% in the control saying they will increase contact from the father while still 41% in the Treatment group ($p = 0.0076$). Furthermore, we can run a regression which controls for whether the respondent knew the information about either study beforehand, whether the experiment was at the beginning or end of the survey, and whether respondents saw one of the slight variations on our messages explained in footnote 19. This indicates that showing information about our findings causes a 13.3 percentage point increase in the likelihood of increasing the proportion of contact from the father ($p = 0.015$).

For those who saw our Treatment message and said they did not know about our study already, 57% reported making no change to their proportion of contact from the father. The most popular reason stated for no change in reaction from a list we allowed them to choose from was because their household already decided that one person should handle all contact with the school. Of those who said a single parent was in charge of all contact with the school 77% reported this to be the mother while the remaining 23% reported it was the father.

Table M.2: Hypothetical Choice Experiment

Control: No information about our study's findings about calls being pushed to fathers	**Treatment**: Information about our study's findings
A recent study finds that when parents become more involved in their children's school lives by receiving regular information about their child's academic progress, their children develop more positive behavior in school. The messages that were used to communicate with the school were sent through a low-cost electronic technology.	A recent study finds that when parents become more involved in their children's school lives by receiving regular information about their child's academic progress, their children develop more positive behavior in school. Another recent study finds that if the child's father emails the school and indicates that he has a lot of availability, about 90% of responses from the school will be directed to the father.

Notes: This table presents the text shown in a between-subjects hypothetical choice experiment.

N Callback Rate

We received a 20% callback rate from our principals, which is in line with prior work. In studies with a similar pool of school principals, the response rate by phone is lower than the response rate via email. Another related outcome is whether principals take a survey in response to an email request, where recent work finds only 14% of principals take this action (Neal et al., 2020). Although not as closely related, recent studies where job applicants submit applications with a phone number and email to employers find that response rates range from 8–11% (Agan and Starr, 2018) to 24% (Kline et al., 2022).

We expected the callback response rate to be lower than the email response rate in previous studies for several reasons. First, making a phone call in response to an email is more time-consuming and takes more effort than hitting "reply." Furthermore, most of the previous studies over-sampled specific types of schools (e.g. charter schools, pre-schools), which makes direct comparisons even more difficult. Bergman and McFarlin Jr (2018) emailed about 6,000 schools with an emphasis on charter schools in 29 states and Washington DC in 2014 and 2018 and find 53% of schools receiving their baseline email respond back via *email*, although 3% of those are automated email responses. Oberfield and Incantalupo (2021) also over-sample charter schools and sent emails to about 3,500 schools in 2018 with an *email* response rate of about 58%. Note that our sample is only about 6% of charter schools. Diaz-Serrano and Meix-Llop (2016) sent emails in Catalonia Spain in 2013 and found about a 60% *email* reply rate. Ahmed et al. (2021) emailed Swedish pre-schools in late 2019 and early 2020 and found their baseline response rate to their emails was 49%. Note, our sample is less than 1% pre-schools. Hermes et al. (2023) emailed about 9,000 childcare centers in Germany in 2021 with about a 71% email response rate. Cantet et al. (2024) emailed about 450 private schools in Columbia in 2022, with 53% receiving an email response. In our study only about 16% of our sample are private schools.

References

Agan, A. and S. Starr (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Agan, A. Y., B. Cowgill, and L. K. Gee (2023). The tradeoffs of transparency: Measuring inequality when subjects know they are in an experiment. *Working Paper*.

Ahmed, A., M. Hammarstedt, and K. Karlsson (2021). Do schools discriminate against children with disabilities? a field experiment in sweden. *Education Economics 29*(1), 3–16.

Bergman, P. and I. McFarlin Jr (2018). Education for all? a nationwide audit study of school choice. *Working Paper*.

BLS (2021). Bureau of labor statistics.

Cantet, N., B. Feld, and M. Hernández (2024). Is there discrimination against children of same-sex households? evidence from an experimental study in colombia. *Labour Economics 87*, 102507.

CBS (2019). Average age of first-time mothers up to 29.9 years. ”https://www.cbs.nl/en-gb/news/2019/19/average-age-of-first-time-mothers-up-to-29-9-years”.

Census (2021). Census bureau survey explores sexual orientation and gender identity. ”https://www.census.gov/library/stories/2021/11/census-bureau-survey-explores-sexual-orientation-and-gender-identity.html”.

Census (2022). Most kids with parent in same-sex relationship live with female couple. ”https://www.census.gov/library/stories/2022/07/most-kids-with-parent-in-same-sex-relationship-live-with-female-couple.html”.

Charness, G., A. Samek, and J. van de Ven (2022). What is considered deception in experimental economics? *Experimental Economics 25*(2), 385–412.

Clark, D. L., L. S. Lotto, and M. M. McCarthy (1980). Factors associated with success in urban elementary schools. *The Phi Delta Kappan 61*(7), 467–470.

Conger, R. D., K. J. Conger, and M. J. Martin (2010). Socioeconomic status, family processes, and individual development. *Journal of marriage and family 72*(3), 685–704.

Daly, M. and F. Groes (2017). Who takes the child to the doctor? mom, pretty much all of the time. *Applied Economics Letters 24*(17), 1267–1276.

Diaz-Serrano, L. and E. Meix-Llop (2016). Do schools discriminate against homosexual parents? evidence from a randomized correspondence experiment. *Economics of Education Review 53*, 133–142.

Hermes, H., P. Lergetporer, F. Peter, F. Mierisch, and S. Wiederhold (2023). Males should mail? gender discrimination in access to childcare. In *AEA Papers and Proceedings*, Volume 113, pp. 427–431. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Kerwin, C., J. Guryan, and J. Pan (2022). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. *Journal of Human Resources*.

Kline, P., E. K. Rose, and C. R. Walters (2022, 06). Systemic Discrimination Among Large U.S. Employers*. *The Quarterly Journal of Economics 137*(4), 1963–2036.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics 3*(4), 303–328.

Nakata, K. (2023). Income, father's and mother's time investment and children's achievement. *Working Paper 0*(0).

NCES (2021). National center for education statistics.

Neal, Z., J. W. Neal, and A. Piteo (2020). Call me maybe: using incentives and follow-ups to increase principals' survey response rates. *Journal of Research on Educational Effectiveness 13*(4), 784–793.

Oberfield, Z. W. and M. B. Incantalupo (2021). Racial discrimination and street-level managers: Performance, publicness, and group bias. *Public Administration Review 81*(6), 1055–1070.

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The ANNALS of the American Academy of Political and Social Science 609*(1), 104–133.

Perez-Vaisvidovsky, N., A. Halpern, R. Mizrahi, and Z. Atalla (2023). Fathers are very important, but they aren't our contact persons: The primary contact person assumption and the absence of fathers in social work interventions. *Families in Society 0*(0), 10443894221145751.

Pleck, J. H. (2007). Why could father involvement benefit children? theoretical perspectives. *Applied development science 11*(4), 196–202.

SSA (2022a). 1980s names. "https://www.ssa.gov/OACT/babynames/decades/names1980s.htm".

SSA (2022b). Last names. "https://namecensus.com/last-names/".

Tzioumis, K. (2018). Demographic aspects of first names. *Scientific data 5*(1), 1–9.

Zamarro, G. and M. Prados (2021). Gender differences in couples' division of childcare, work and mental health during covid-19. *Review of Economics of the Household 19*(1), 11–40.