Who you gonna call?: Gender inequality in demand for parental involvement*

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

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

January 25, 2023

Abstract

Prior studies find a significant inequality in time spent by men and women in heterosexual households on child-related tasks even when both work full time. This inequality has been linked to gender inequalities in a wide range of labor market outcomes, human capital accumulation, and economic mobility. We investigate an important potential source of this inequality: external demands for parental involvement. We pair a novel theoretical model with a large-scale field experiment and find that mothers are 1.4 times more likely to be contacted than fathers by their child's school. We decompose this inequality into discrimination stemming from differential beliefs about parents' availability and other factors to better inform policies aimed at closing gender gaps. Moreover, variation in demographic characteristics across the universe of schools and principals allows us to study the implications of our findings for different types of principals, schools and households.

JEL Classification:
Keywords: Discrimination, Gender Gap, Field Experiment

*We thank Katelyn Cranney, Paris Rich, and Katherine Jolley for their excellent research assistance and Fran Blau, Amanda Chuan, Zoe Cullen, Ingrid Haegele, Louis-Pierre Lepage, Betsey Stevenson, and Basit Zafar for their valuable feedback on earlier drafts of the paper. Special thanks to Francis Ditraglia for many useful conversations. We express our appreciation to audience members at the SITE-Gender, the Economic Science Association Conference, the Discrimination and Disparities Seminar, the BYU/University of Utah Applied Micro workshop, Behavioral Experimental Design Initiative, Stockholm University (SOFI), Uppsala University, Gothernburg University, Middlebury College and the Association for Social Sciences (ASSA) meetings. We thank the Brigham Young University College of Family, Home, and Social Science and Tufts University for their generous research support. This work has been supported (in part) by Grant 2204-38170 from the Russell Sage Foundation. Any opinions expressed are those of the principal investigators alone and should not be construed as representing the opinions of the funder. This study was pre-registered at the AEA registry (AEARCTR-0007610) and approved by the relevant IRBs.

1 Introduction

Despite the convergence of men and women's roles in the labor market, there remains a persistent gender earnings gap. Recent statistics in the U.S. show that women employed full time earn about 18 percent less than their male counterparts (US Census Bureau, 2020). Extensive prior literature finds that an important contributor to this gap is women's tendency to concentrate in occupations with more temporal flexibility.[1] In both observational and experimental work, women are more likely than men to value flexible work arrangements including part-time work, flexible schedules, and working from home. This is especially true for women with children.[2]

The need for greater workplace flexibility is consistent with the robust finding that women engage in a disproportionate share of child-related tasks relative to men.[3] U.S. time-use data shows that married mothers employed full time spend over 50 percent more time caring for children than analogous fathers, even though couples' support for a more equal division of household labor has risen considerably in recent years (BLS, 2021; Scarborough, Sin, and Risman, 2019). In two-parent heterosexual households with both parents working full-time, 35 percent of mothers report experiencing a household interruption during their workday, compared to only 20 percent of fathers (Cubas, Juhn, and Silos, 2021). These imbalances have been linked to substantial gender inequalities in a wide range of economic measures, including labor market outcomes, human capital accumulation, and economic growth.[4]

In this paper, we investigate a previously unexplored source of this inequality: external demands on parents' time. We develop a theoretical model to inform the design of a field experiment in a K-12 school setting. Specifically, we send emails with phone numbers for both parents in a two parent household to the near-universe of U.S. school principals (N = 80,071) asking the principal to contact a parent by phone about a general school-related inquiry. We randomly vary the information about parents' availability to disentangle whether discrimination stems from decision-maker

[1] Price and Wasserman (2022); Wasserman (2022); Goldin (2014); Flabbi and Moro (2012); Goldin and Katz (2011)

[2] Price and Wasserman (2022); Wiswall and Zafar (2018); Mas and Pallais (2017); Goldin and Katz (2011); Duchini and Van Effenterre (2022)

[3] Aguiar and Hurst (2007); Craig and Mullan (2011); Schoonbroodt (2018).

[4] Adams-Prassl (2023); Erosa, Fuster, Kambourov, and Rogerson (2022); Cubas, Juhn, and Silos (2022); Albanese, Nieto, and Tatsiramos (2022); Cubas et al. (2021); Duchini and Van Effenterre (2022); Kuziemko, Pan, Shen, and Washington (2018); Kleven, Landais, and Søgaard (2019); Angelov, Johansson, and Lindahl (2016).

beliefs about parents' responsiveness (for example, that women are more available because they are stay-at-home-moms or because they want to be more involved in the decision) or other deterrents (for example, distaste, systemic factors, social norms or beliefs not related to responsiveness). This allows us to investigate whether the gender gap can be mitigated by encouraging households to change the signals they send. Furthermore, we explore whether attributes of the external decision maker impact the inequality in demands on parents' time, in which case the gender gap might be mitigated through a policy targeting behavioral change in certain types of decision-makers or schools.[5]

We document striking gender and treatment differences. Principals are significantly more likely to call mothers first in the baseline treatment which contains no signal about parents' availability. On average, mothers are called first 1.4 times more than fathers (58 vs. 42 percent). This provides direct and novel evidence of greater external demands on mothers in our setting. We believe our findings are a first step towards documenting gender inequality in external demands which are likely present within the school setting but in a wider range of tasks (e.g. picking up a sick child, volunteering for school events) as well as outside the school setting (e.g. who schedules doctor visits, coordinates extracurricular activities, or who grandparents expect to take care of a child's needs). Thus, our work likely represents a lower-bound on the overall demands on mothers' versus fathers' time.

In addition to documenting this important gender gap, we explore the reasons it arises and test potential solutions. Specifically, we show that signaling that the father is more available mitigates the inequality and causes mothers to be called less than half the time. It is notable, however, that even when fathers signal that they are more available, mothers still get 26 percent of the calls. In contrast, signals that reinforce stereotypes that fathers are less available and mothers are more available cause mothers to receive between 73-90 percent of calls. Strikingly, even when the email comes from the father *and* the father signals his availability, 12 percent of the calls are still directed to mothers. This highlights an important asymmetry in the effectiveness of informational interventions in closing the observed gender gap in external demands for parents' time.

[5]The scope of this paper is to think exclusively about two-parent households with a male and female parent. However, we acknowledge that there are many types of households and more than the male and female gender. We believe exploring the effect of external demands in other settings is an important question for future work and discuss this in Section B.

We pair a novel theoretical model with our field experiment in order to identify the mechanisms underlying any differential demand for parental involvement. Our random utility model shows how decision makers decide whether to contact a mother, father, or neither parent as a function of the expected benefit of a response, the cost of making contact, and other deterrents. The model allows us to attribute any differences we find in the rate of calls to mothers versus fathers to either decision-makers' beliefs about parents' responsiveness or other factors (which could be a result of distaste, systemic factors, or beliefs unrelated to responsiveness).

Using this modeling framework, we measure the impact of beliefs about expected responsiveness by randomizing the signals we send to decision makers about the availability and/or involvement of a specific parent. The proportions of calls to mothers, fathers, or neither parent under the various signals are what allow us to decompose the reasons for gender inequality when no signals are present. The logic is as follows. The signal impacts only decision-maker's beliefs about the benefit of a call, so any differences in the proportions of calls to a parent across signals tell us what happens when those beliefs change. Once we have identified the beliefs about responsiveness, the residual differences in the proportions of calls to mothers versus fathers reveal the extent to which gender differences are driven by other deterrents. Understanding the reasons for the gender inequality informs which policies can be used to mitigate the gender gap.

The model is flexible enough that we can use it to explore various belief-based channels, such as parental availability versus involvement or expertise, which we do in a separate treatment. We find that signaling that both parents want to be equally involved in the decision does not reduce the share of calls to mothers suggesting that beliefs about parents' relative expertise or involvement are not the primary driver of the results. The model also allows decision-maker beliefs as well as other deterrents to vary by decision-maker and school characteristics.

This paper extends the existing literature in four important ways. First, we empirically document a novel gender gap in external demand for parental involvement. While prior research has found that women spend significantly more time on child-related tasks than men in similar households, our study is the first to explore if this inequality is partially driven by external demands for parental involvement and to experimentally test potential solutions. The novel gender inequality that we document has significant and persistent economic consequences for women and men, who

report a desire for a more equal distribution of child-related tasks (Pew Research, 2015). In our own survey of parents in heterosexual households with school-age children, which we detail in section F, we find that women report being contacted by the school more often than men; yet women wish they were contacted less often, while men wish they were contacted more often. Greater external demands for parental involvement may push women towards more flexible jobs leading to substantial labor market penalties, including labor force participation, occupational choice, and earnings.[6] To quantify the economic consequences of this inequality, Cubas et al. (2021) use the timing of labor supplied during the day to document that full-time employed mothers experience more work-day interruptions than their male counterparts and that these interruptions reduce wages by about 9 percent.

Relatedly, prior research has documented the effects of childcare disruptions on women's labor market outcomes. The Covid-19 pandemic and the associated school and daycare closures, for example, led to significantly larger declines in women's employment and labor force participation relative to men. The negative effects have been especially large for mothers of school-aged children, leading to significant declines in their mental and physical health.[7]. Similarly, Wasserman (2022) documents a significant decline in women's labor market activity during summer months and provides compelling evidence that summer childcare constraints contribute to career choices and earnings, particularly for women with school-aged children in line with findings from Duchini and Van Effenterre (2022). Understanding whether external demands for parental involvement contribute to gender inequalities in child-related tasks can shed light on the drivers of the persistent gender earnings gap and inform policies aimed at mitigating persistent gender inequalities.

Secondly, we contribute to the growing literature on the role of information provision in reducing discrimination. Prior work in economics and social psychology has considered the role of individual-specific information in reducing reliance on group statistics for evaluations (also known as statistical or belief-based discrimination). This literature has produced mixed evidence. While several recent studies show

[6] Adda, Dustmann, and Stevens (2017); Cortes and Pan (2016); Pertold-Gebicka, Pertold, and Datta Gupta (2016); Goldin (2014); Gicheva (2013).

[7] Couch, Fairlie, and Xu (2022); Garcia and Cowan (2022); Hansen, Sabia, and Schaller (2022); Amuedo-Dorantes, Marcén, Morales, and Sevilla (2020); Zamarro and Prados (2021); Sevilla and Smith (2020); Montes, Smith, and Leigh (2021); Heggeness (2020); Russell and Sun (2020); APA (2021)

that providing accurate information reduces statistical discrimination (Laouénan and Rathelot, 2022; Bohren, Imas, and Rosenberg, 2019); others have found no discernible effects (Bertrand and Mullainathan, 2004; Oreopoulos, 2011).

Our paper advances this literature by documenting a striking asymmetry in the effect of information on reducing discrimination. In our field experiment, we test whether providing information about parents' availability mitigates the gender gap in external demands for parental involvement. Notably, while we find that signaling father's availability moves calls away from mothers, we also document the limits of this informational intervention. Specifically, we find that signaling mother's high availability leads to mothers being contacted 90 percent of the time, while signaling father's high availability increases calls to fathers only up to 74 percent.

A related literature to which we contribute investigates the underlying sources of discrimination. While field experiments lend themselves to identifying the existence of discrimination and its incidence, few are able to identify the mechanisms that lead to discriminatory behavior (Bertrand and Duflo, 2017). The two most-studied mechanisms for discrimination in economics are tastes/preferences (Becker, 1957) and beliefs (Phelps, 1972; Arrow, 1973; Aigner and Cain, 1977), with recent work emphasizing the importance of indirect discrimination stemming from systemic and institutional factors (Bohren, Hull, and Imas, 2022; Kline, Rose, and Walters, 2022). We join a small, but growing, literature that attempts to differentiate these sources of discriminatory behavior.

While differentiating these sources is possible in observational data only under special circumstances (e.g., Knowles, Persico, and Todd, 2001; Sarsons, 2017), field experiments are a more fertile ground in this regard. List (2004) combines information from several sources, including a field experiment, to conclude that racial discrimination in the baseball card market is statistical and not preference-based. Islam, Pakrashi, Wang, and Zenou (2018) uses a simple model combined with institutional details to quantify the proportion of patients who employ statistical discrimination when choosing a physician. Bohren et al. (2019) combine a theoretical model with dynamic data and variation in the subjectivity of decisions to conclude that gender discrimination on a mathematics Q&A forum is belief-based, with a component that comes from incorrect beliefs. Nunley, Pugh, Romero, and Seals (2015) use the levels-versus-variance approach of Neumark (2012) to find taste-based discrimination as the most likely source of racial discrimination in a job-market correspondence study. We

advance this literature by using a simple, static theoretical model combined with a field experiment to identify separate parameters that capture the beliefs and preferences underlying discriminatory behavior.

Finally, this paper contributes to the rich literature on institutional (also known as structural or systemic) discrimination. Prior work in sociology and economics has explored the idea that discrimination may be perpetuated by organizations or structures, in addition to individuals (for discussions, see Small and Pager (2020); Bohren et al. (2022); Kline et al. (2022); Scott (2013); Council (2004); Powell and DiMaggio (2012)). We provide novel evidence of institutional discrimination in our setting, where school principals aren't merely individual agents but represent powerful organizational norms and practices. As Small and Pager (2020) argue, institutional discrimination deserves particular attention given the stability and deeply ingrained nature of systemic practices and their long-lasting consequences.

Notably, the patterns that we document represent only a small share of the overall gender inequality in external demands for parental involvement. While the gender gap in school-related interruptions closely mirrors gender gaps in other child-related and household domains, this is only one of many settings where mothers are disproportionately more likely to experience child-related interruptions on a daily basis.[8]. The gender inequality in physical housework, for example, has remained largely unchanged since the mid-1990's, with men spending about half as much time on housework as women in similar households (Bianchi, Sayer, Milkie, and Robinson, 2012). Furthermore, men's housework hours tend to be disproportionately allocated toward relatively infrequent and flexible tasks (e.g., home repairs or yard work), while women shoulder many of the recurring daily tasks (e.g., cooking and childcare) that cannot be put off to a convenient time (Bianchi et al., 2006). Moreover, research across social sciences has increasingly drawn attention to "invisible" forms of labor, including emotional and cognitive labor being disproportionately shouldered by women.[9] While these inequalities are more difficult to measure directly, our findings shed light on potential policies aimed at mitigating these gender gaps.

[8]See for example, Wikle and Cullen (2021); Bianchi, Robinson, and Milke (2006); Boye (2015); Daly and Groes (2017); Daminger (2019); Bertrand, Kamenica, and Pan (2015b); Charmes (2019)

[9]Daminger (2019); Offer (2014); Lee and Waite (2005).

2 Theoretical Framework

In this section we present our theoretical framework. Our aim is to show how a decision-maker who interacts with a two-person, heterosexual couple decides which person to call upon for a task. In our specific field experiment, the decision-maker is a school principal and the task is a discussion about enrolling at the school. However, our theoretical model is flexible enough to be applied to different types of decision makers (e.g. doctors, dentists, school teachers, sports coaches, music teachers, summer camp directors, or organized religion leaders) and different types of tasks (e.g. picking up a sick child, waiting in line to enroll in lessons, volunteering for career day or a bake sale, taking the team on an overnight trip). Furthermore, our model could also apply to differences in demands on parents but also about any type of demand on a two-person household (e.g. for elder care, interior design projects, home renovations, retirement planning).

We lay out a very simple economic structure in Section 2.1 to capture the decision-making behavior of school principals when contacting parents. In Section 2.2, we describe the random utility model we use to study this environment. We then explain in Section 2.3 how our experimental variation integrates with the random utility model. Section 2.4 shows how we use the model to identify and estimate the structural parameters of the model, most notably the beliefs and taste parameters for the principal.

The model is quite flexible and can be extended in several directions. In Section 2.5, we discuss robustness and add two extensions. One captures heterogeneity in the characteristics of principals and the other shows that the model can also easily accommodate interdependence between principals' beliefs about male and female parents.

2.1 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 $j = 1, \ldots, J$ or at our exogenously-determined experiment end date. The observables in our experiment are then (1) the choice $y_i \in J$ for each decision maker and (2) the characteristics of the alternatives

x_i that are shown to each decision maker.

We assume that decision makers potentially face different costs, c_i, of making a phone call. For instance, some may have inferior technology or be busier than others. We also assume that decision makers potentially perceive different benefits from choosing different alternatives. We further assume that there are two components to these benefits: the decision maker's belief about how likely an alternative is to respond to their call and the decision maker's distaste for calling that alternative. We let r_{ij} denote decision maker i's subjective probability assessment that alternative j will respond[10] and we let δ_{ij} denote i's "distaste" for calling alternative j.

We assume that each decision maker i knows c_i, δ_{ij}, has beliefs over r_{ij} and is risk neutral.

2.2 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 expected utility maximizer i, the expected utility of calling alternative j is defined as

$$U_{ij} = \mathbb{E}(r_{ij}) - \delta_{ij} - c_i. \tag{1}$$

where δ_{ij} is positive if i has a distaste for calling alternative j and negative if i has a taste for calling alternative j. This is our basic random utility formulation.

Because calling no one incurs no cost and provides no benefit, we take the utility of *call neither* to be zero. This normalization will play a crucial role in identification because choice in this context is determined by differences in utility, not levels.

Under this normalization and in our context of the choice between two parents and calling neither, decision maker i calls neither parent if both $U_{i,m} < 0$ and $U_{i,f} < 0$; calls the female parent if $U_{i,f} \geq 0$ and $U_{i,m} \leq U_{i,f}$; and calls the male parent if $U_{i,m} \geq 0$ and $U_{i,f} < U_{i,m}$.[11]

We can break the decision maker's choice into two parts. First, they decide whether to call any parent. If they decide to call someone, then they must decide which parent to call.

[10] We assume that the alternative who is called will have the relevant expertise with probability 1. It is straightforward to add a non-degenerate belief about expertise, which we do in Section 2.5.1.

[11] We break ties in favor of calling the female parent, although utility is continuous so this has no impact in terms of the theory.

The decision maker makes a call if and only if

$$\max\{\mathbb{E}(r_{i,f}) - \delta_{i,f},\, \mathbb{E}(r_{i,m}) - \delta_{i,m}\} \geq c_i$$

If the decision maker decides to make a call, they call the female parent when

$$\mathbb{E}(r_{i,f} - r_{i,m}) \geq (\delta_{i,f} - \delta_{i,m})$$

and they call the male parent when

$$\mathbb{E}(r_{i,m} - r_{i,f}) < (\delta_{i,m} - \delta_{i,f}).$$

Notice that the cost, c_i, does not affect the decision between parents 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, whereas the choice of which parent to call depends only on the differences in beliefs and distaste across parents.

2.3 Experimental Manipulation of Beliefs

Consider an experimental manipulation that sends an informative signals to decision maker i about either their belief about the female ($r_{i,f}$) or male parent ($r_{i,m}$). For simplicity, we assume all priors and signals are normally distributed. That is,

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

where $\bar{r}_j$ and ω_j^2 are the prior mean and variance common to all i. x_{ij} is a signal of the *true* responsiveness r_j of j that we send to i and the signal variance is σ^2.

We assume that the priors for $r_{i,f}$ and $r_{i,m}$ are independent of each other and also of the distributions for the cost and distaste parameters. This implies that when we send a signal about one parent (female or male), it shifts only the belief about the parent for which the signal was sent and, further, does not impact the δ_{ij}'s or c_i.[12] Because of our assumption that decision makers are risk neutral, only the marginal means of this distribution are relevant for expected utility and therefore decisions.

We then have that decision maker i's posterior mean $\widetilde{r}_{ij}$ for the responsiveness of

[12] We relax this assumption is Section 2.5.2 below.

parent j is

$$\widetilde{r}_{ij} = \lambda_j \bar{r}_j + (1-\lambda_j)x_{ij}, \quad \lambda_j = \frac{1/\omega_j^2}{1/\omega_j^2 + 1/\sigma^2} \tag{2}$$

Letting w_{ij} be an indicator for sending i a signal regarding r_j and recalling that x_{ij} is the value of the signal, Equation 1 becomes

$$\begin{aligned} U_{ij} &= (1-w_{ij})\bar{r}_j + w_{ij}\widetilde{r}_{ij}(x_{ij}) - \delta_{ij} - c_i \\ &= (1-w_{ij})\bar{r}_j + w_{ij}\left[\lambda_j \bar{r}_j + (1-\lambda_j)x_{ij}\right] - \delta_{ij} - c_i \\ &= \bar{r}_j - (1-\lambda_j)\bar{r}_j w_{ij} + (1-\lambda_j)w_{ij}x_{ij} - \delta_{ij} - c_i. \end{aligned} \tag{3}$$

for $j \in \{f, m\}$. Recall that the utility of calling neither parent ($U_{i,n}$) is assumed to be zero.

Using $\bar{\delta}_j$ to denote the average value of δ_{ij} and c to denote the average value of c_i across the distribution of principals, Equation 3 can be written as

$$U_{ij} = \alpha_j + \eta_j w_{ij} + \gamma_j w_{ij} x_{ij} + \varepsilon_{ij} \tag{4}$$

$$\alpha_j = \bar{r}_j - \bar{\delta}_j - c \tag{5}$$

$$\eta_j = -(1-\lambda_j)\bar{r}_j \tag{6}$$

$$\gamma_j = 1 - \lambda_j \tag{7}$$

$$\varepsilon_{ij} = (c - c_i) + (\bar{\delta}_j - \delta_{ij}). \tag{8}$$

The ε_{ij} are econometric errors. They are mean zero because the average terms $\bar{\delta}_j$ and c are absorbed in the constant α_j. Importantly, the random assignment of x_{ij} and w_{ij} imply that they are independent of ε_{ij}.

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

2.4 Identification of Reduced-Form and Structural Parameters

Identification is straightforward given the following elements of our setting and our model:

1. The random utility model provides structure for the relationship between benefits, costs and outcomes.

2. The "call neither" outcome provides a clear normalization because it provides no benefits and incurs no costs.

3. Experimental randomization establishes that the regressors are not dependent on the outcome variable.

4. The assumption that errors are drawn from the logistic distribution leads to equations for the outcome probabilities that are closed-form.

This would be a standard random utility model if our reduced-form parameters α_j, γ_j, and η_j did not vary across the j choices. Having intercepts and slopes that vary across alternatives is, however, crucial to learning about how the experimental manipulation impacts the choices of decision makers. Fortunately, the structure of the model allows us to identify these intercepts and slopes.

Here we state the identification result and provide intuition for this result. All proofs are in Appendix G, with the proof from this section in Appendix G.2.

Result 1. *Given the assumptions of Sections 2.1–2.3, the reduced-form parameters α_j, γ_j, and η_j are identified for $j \in \{f, m\}$.*

We identify the reduced-form parameters using ratios of the proportions of signal-outcome pairs. We denote the proportions as p_j^t where subscripts indicate alternative and superscripts indicate treatment, where $t \in \{b, l\boldsymbol{F}, h\boldsymbol{F}, l\boldsymbol{M}, h\boldsymbol{M}\}$ where b indicates the baseline 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. For example, $p_m^{l\boldsymbol{F}}$ is the proportion of decision makers who receive the low signal about female parent availability and then call the male parent.

Given the assumption that $\alpha_n = 0$, each α_j intercept is directly identified by comparing the proportions of decision makers who receive no signal and call parent j (p_j^b for $j \in \{f, m\}$) and the proportion of decision makers who receive no signal and call neither parent (p_n^b).

To separately identify γ_j and η_j, we need to create variation in the term $w_{ij}x_{ij}$, i.e., the interaction of the indicator variable for whether a signal was sent (w_{ij}) and

the value of the signal (x_{ij}). This variation needs to be distinct from the variation in w_{ij} alone. We achieve this by sending two values of the signal about each alternative j with cardinal values that are known. Specifically, we send both a positive signal and a negative signal about each parent and assume the values are 1 and -1.[13]

We now turn to the identification of the structural parameters $\bar{r}_j$, $\bar{\delta}_j$, c, and λ_j, where we have the following result:

Result 2. *Given the assumptions of Sections 2.1–2.3 and Result 1, the structural parameters $\lambda_f, \lambda_m, \bar{r}_f, \bar{r}_m$ and $\bar{\delta}_m - \bar{\delta}_f$ are identified.*

Since we can identify γ_j, Equation 7 provides for identification of λ_j, which is the weight that decision makers place on their prior belief when updating. Given η_j and λ_j, Equation 6 allows us to identify the prior belief $\bar{r}_j$. Combining $\bar{r}_j$ and α_j, Equation 5 identifies $\bar{\delta}_j + c$. We can then combine $\bar{\delta}_f + c$ and $\bar{\delta}_m + c$ to identify the difference in beliefs about female and male parents, $\bar{\delta}_f - \bar{\delta}_m$.

We can develop intuition by looking at the the relationships between the reduced-form and structural parameters.

$$\begin{aligned}
\bar{r}_j &= -\frac{\eta_j}{\gamma_j} \\
\bar{\delta}_m - \bar{\delta}_f &= \frac{\eta_f}{\gamma_f} - \frac{\eta_m}{\gamma_m} + \alpha_f - \alpha_m \\
\lambda_j &= 1 - \gamma_j
\end{aligned}$$

The second equation is derived from the definition of α_j and can be expressed as

$$\alpha_f - \alpha_m = \bar{\delta}_m - \bar{\delta}_f + \bar{r}_f - \bar{r}_m \tag{9}$$

We can interpret this as indicating that the magnitude of the motherhood effect (if indeed $\alpha_f > \alpha_m$) derives from the excess distaste decision makers have for calling male parents plus their excess belief in the availability of female parents.

With both the reduced-form and structural parameters identified, we can now state the following testable implications of the theory:

Hypothesis 1. *There is a "motherhood effect." In order to detect a motherhood effect, we need to see $p_f^b > p_m^b$, that is, the proportion of decision makers who receive*

[13]For a discussion of the impact of the chosen scale of signals, see Section 2.5.4.

no signal and call the female parent is larger than the the proportion who receive no signal and call the male parent. This is equivalent to $\alpha_f > \alpha_m$ in terms of the reduced-form parameters and $\bar{r}_f - \bar{\delta}_f > \bar{r}_m - \bar{\delta}_m$ in terms of the structural parameters.

Hypothesis 2. *Without intervention, decision-makers believe that female parents are more available than male parents. We find support for this hypothesis if $\bar{r}_f > \bar{r}_m$.*

Hypothesis 3. *On average, decision makers' distaste for calling male parents is larger than their distaste for calling female parents. We find support for this hypothesis if $\bar{\delta}_m - \bar{\delta}_f > 0$.*

We can combine the testable implications in Hypotheses 1–3 with Equation (9) to say the following about the sources of differential treatment of female and male parents.

Result 3 (Sources of differential treatment)**.** *If there is a "motherhood effect" [Hypothesis 1] and*

1. *Hypothesis 2 is supported (i.e., $\bar{r}_f > \bar{r}_m$) while Hypothesis 3 is not supported (i.e., $\bar{\delta}_m - \bar{\delta}_f \leq 0$), then differential treatment is beliefs-based.*
2. *Hypothesis 3 is supported (i.e., $\bar{\delta}_m - \bar{\delta}_f > 0$) while Hypothesis 2 is not supported (i.e., $\bar{r}_f \leq \bar{r}_m$), then differential treatment is preference-based.*
3. *both Hypothesis 2 (i.e., $\bar{r}_f > \bar{r}_m$) and Hypothesis 3 (i.e., $\bar{\delta}_m - \bar{\delta}_f > 0$) are supported, then differential treatment is based in both beliefs and preferences.*

The parameters required to test Hypotheses 1–3 are identified from data as demonstrated in Result 2.

2.5 Model Extensions and Robustness

2.5.1 Beliefs about both availability and expertise

Until now, we have assumed that decision-maker beliefs about the value of calling parents only incorporate the parents' availability. We can expand our conceptualization of decision-maker beliefs to also include the expertise of the parents. If we model

these two components of beliefs as multiplicative, i.e., $\mathbb{E}[r_j q_j] = \overline{r}_j \overline{q}_j$, decision-maker utility in Equation (1) becomes

$$\mathbb{E}(U_{ij}) = \overline{r}_j \overline{q}_j - \delta_j - c_i.$$

We now re-examine the identification of our structural parameters given this new element of the model. In order to do so, we must take a stand on how decision makers will interpret our signals about availability now that their beliefs also contain the expertise component.

If the signal about availability does not impact the belief about expertise, the prior belief about expertise is simply carried along with the signal, so that the updated belief is

$$\widetilde{qr}_{ij} = \lambda_j \overline{q}_j \overline{r}_j + (1 - \lambda_j) \overline{q}_j x_{ij}$$

and expected utility after updating on the signal is

$$\begin{aligned}\mathbb{E}(U_{ij}) &= (1 - w_{ij}) \bar{q}_j \bar{r}_j + w_{ij} \widetilde{qr}_j(x_{ij}) - (\delta_j + c_i) \\ &= \bar{q}_j \bar{r}_j - (1 - \lambda_j) \bar{q}_j \bar{r}_j w_{ij} + (1 - \lambda_j) \bar{q}_j w_{ij} x_{ij} - (\delta_j + c_i).\end{aligned}$$

The following equations now map the reduced-form parameters to the structural parameters:

$$\begin{aligned}\alpha_j &= \bar{q}_j \bar{r}_j - \bar{\delta}_j - c \\ \eta_j &= -(1 - \lambda_j) \bar{q}_j \bar{r}_j \\ \gamma_j &= (1 - \lambda_j)\, \bar{q}_j\end{aligned}$$

and we have $\eta_j = -\gamma_j \bar{r}_j \;\Leftrightarrow\; \bar{r}_j = -\frac{\eta_j}{\gamma_j}$ as in the base model. That is, our experimental variation continues to identify the availability belief even when the belief contains more than just availability. However, we no longer cleanly identify the updating or distaste parameters. Instead, we have $\lambda_j = \frac{1-\gamma_j}{\bar{q}_j}$ and $\bar{\delta}_j + c = -\bar{q}_j \frac{\eta_j}{\gamma_j} - \alpha_j$. Both are polluted by the average belief about expertise. Moreover, the distortion depends on the both the sign and magnitude of the average belief about expertise.

To address this concern, we introduce a variation on our main set of five treatments (henceforce "Main" variation), which we term the "Equal Decision" variation. In the "Equal Decision" variation, we send the same set of five messages, but to each we add

the statement, "This is the type of decision we both want to be involved in equally" in order to fix the decision-maker's belief about parental expertise.

If we do not take a stand on the value of the expertise signal and label that value $q_j^{'}$, the updated belief becomes

$$q_j^{'}\widetilde{r}_{ij} = \lambda_j q_j^{'}\bar{r}_j + (1-\lambda_j)q_j^{'}x_{ij}$$

All five treatments, including the baseline, receive this same message $q_j^{'}$ about expertise, so it appears in both terms on the right hand side. Similar to the "Main" variation discussed above, we once again are able to identify the beliefs about availability as $\bar{r}_j = -\frac{\eta_j}{\gamma_j}$, but we do not get $\bar{\delta}_j + c$ cleanly; instead we have $\bar{\delta}_j + c = -q_j^{'}\frac{\eta_j}{\gamma_j} - \alpha_j$.

If we are willing to assume that $q_j^{'} = 1$ (the same value we have assumed for positive signals about availability), then we cleanly identify the distaste + cost parameter as the same combination of reduced-form parameters as in the "Main" variation where we assume there is no belief about expertise. In fact, if we are willing to assume any particular value for the expertise signal, we can cleanly identify the distaste + cost parameter in this more general case where beliefs include an expertise component.

We can actually do more to understand beliefs about expertise. If we combine the baseline treatment from the "Main" variation with the remaining four treatments from the "Equal Decision" variation with $q_j^{'} = 1$, we have

$$\widetilde{q}r_{ij} = \lambda_j\bar{q}_j\bar{r}_j + (1-\lambda_j)x_{ij}q_j^{'} = \lambda_j\bar{q}_j\bar{r}_j + (1-\lambda_j)x_{ij}$$

The mapping from reduced form to structural parameters becomes

$$\begin{aligned}
\alpha_j &= \bar{q}_j\bar{r}_j - \bar{\delta}_j - c \\
\eta_j &= -(1-\lambda_j)\bar{q}_j\bar{r}_j \\
\gamma_j &= (1-\lambda_j)\,.
\end{aligned}$$

Similar to the "Main" variation that ignores beliefs about expertise, γ_j identifies λ_j and we can recover both $\bar{q}_j\bar{r}_j = -\frac{\eta_j}{\gamma_j}$ and $\bar{\delta}_j + c = -\frac{\eta_j}{\gamma_j} - \alpha_j$. The only difference is that the belief now encompasses expertise. Given enough statistical power, we can then divide $\bar{q}_j\bar{r}_j$ by the $\bar{r}_j$ identified in the "Main" variation to recover $\bar{q}_j$ separately.

2.5.2 Adding decision-maker characteristics

Until now, we have assumed that all decision makers are identical in terms of their observable characteristics. We can, however, easily allow for decision makers to differ in their beliefs and tastes according to an observable characteristic; we are especially interested in the gender of the principal.

To be clear: we do not change the signals that we send to principals in any way. This model extension simply allows the signals we send to impact the beliefs of different types of decision makers differently. Appendix G.3 contains the details of this model extension, identification result and additional testable hypotheses.

2.5.3 Relaxing the independence assumption

Now suppose that a signal about one parent induces the decision maker to update their belief about *both* parents.[14] This could happen, for instance, if the decision-maker's beliefs about the parents are correlated or the decision maker directly infers information about both parents from a signal about just one parent.

This model extension allows us to identify the magnitude of the impact of signals about one parent on the belief about the other parent. It also allows us to determine whether decision makers put different weight on their prior beliefs when signals are about females versus males. The generalized utility formulation and mapping to reduced-form and structural parameters is available in Appendix G.4, along with the identification result and an additional testable hypothesis.

2.5.4 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 values of the signal 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* (that is, η_j) does not change. However, although the *value* of the signal is now assumed to be different, the term $(1 - \lambda_j)w_{ij}x_{ij}$ in Equation 3 does not vary with

[14] We present the theory for the case where decision-makers are not differentiated by characteristic as in Section 2.4. It is straightforward to combine the two extensions to have both decision-maker characteristics and cross impact of signals.

our assumption about the value of x_{ij}. Instead, when we change x_{ij}, $\gamma_j = (1 - \lambda_j)$ 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, as they are entirely determined by the baseline.

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.[15] 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.

2.5.5 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 is actually 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. Appendix G.5 provides additional intuition for the case of risk-averse decision makers.

[15] Note that the treatment effects parameters would *not* change as long as they are still just dummy variables for each treatment. The simple relationships between the treatment effects and the reduced-parameters would be modified to take account of the fact that the treatment effects model ignores the value of the signals.

3 Field Experiment

Our theoretical model informs the design of a large-scale field experiment. The experiment consists of sending email messages to a near-universe of U.S. school principals from a fictitious couple with one male parent and one female parent.[16] Email is a common way for parents to contact schools. Our own survey of educators found that three-fourths of them report being contacted by parents via email at least once a month[17]. Additionally, several recent studies have used emailing schools as part of their methodology to document discrimination against students with disabilities, of certain races, or with homosexual parents (Oberfield and Incantalupo, 2021; Bergman and McFarlin Jr, 2018; Ahmed, Hammarstedt, and Karlsson, 2020).

3.1 Setting

Our experiment takes place in a K-12 school setting. A large portion of the general population, about 40% of households in the US, have school-aged children (NCES, 2021). Schools are an ideal setting to explore external demands on parents' time because 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, volunteering for the book fair or joining the Parent Teacher Association for several reasons. First, educators in our survey report that they would favor contacting the mother first in many of these scenarios (we discuss the survey in Section F). Second, the gender distribution of these tasks is significantly skewed with mothers comprising almost 90% of Parent Teacher Association members and only 13% of fathers reporting high levels of involvement in their child's school activities, compared to 53% of mothers.[18]

[16] We acknowledge that there are many different types of households, and more than two genders and discuss this further in Section B. We describe our data collection process in more detail in Section E as well as some of the ethical considerations in section C.

[17] We discuss the survey in detail in Section F.

[18] 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

Furthermore, although the gender inequality that we document is in the school setting, this is only one of many settings where mothers spend significantly more time on child-related tasks than fathers. Prior studies have documented substantial gender differences in time devoted to caring for sick children, taking children to the doctor, and coordinating a wide range of household and child-related tasks, known as cognitive labor.[19] If these other inequalities are also partially driven by external demands, our findings likely represent a lower bound for the overall gender gaps in external demand for parental involvement.

3.2 Messages

In our experiment, school principals receive an email from a two-parent heterosexual household. The email states that the parents are searching for a school for their child and would like to have a phone discussion about it. We provide separate phone numbers for each parent and randomize the order of the phone numbers listed as well as whether the sender is male or female.[20] We developed the specific message in consultation with school administrators from a variety of schools (public, private and charter schools). Our conversations and survey evidence (Appendix F) confirmed that parents frequently make general email inquiries to schools prior to enrolling and that it is common for one parent to email, copying the other parent.

We then augment our baseline message across our other treatments by adding a sentence indicating the availability of a specific parent in the two-parent household. We show examples of the exact variation in wording in Figure 1. Details of the exact names and email addresses used in the experiment are in Appendix E and the full text of messages is in Appendix I.

We designed these messages based on our theoretical model discussed in Section 2. One of the key results of the model is that by varying the strength (low/high) of the signals about each of our parents' availability, we can disentangle the extent to which the gender inequality is driven by beliefs about mothers having higher availability versus other deterrents.

[19] Wikle and Cullen (2021); Bianchi et al. (2006); Boye (2015); Daly and Groes (2017); Daminger (2019); Bertrand et al. (2015b); Charmes (2019) `https://www.bls.gov/spotlight/2017/differences-in-parents-time-use-between-the-summer-and-the-school-year/home.htm` `https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0009/3963249/HILDA-Statistical-Report-2021.pdf`

[20] The sender's phone number is always listed first.

3.3 Sample Frames

During the summer of 2022 we sent emails to a near-universe (a sample of 80,071) of school principals across the U.S. We begin by describing the "Main" variation of our experiment which was sent to 30,471 school principals.

3.4 Data Collection

We observe whether any call is made to any of the phone numbers we list including phone calls where no voicemail was left. We also know the precise time, date, and content/length of any voicemails left for our parents. We use this information to match each phone call back to the original decision-maker who received one of our treatment emails. Appendix E provides more details about the experimental design, data collection and matching process.

Approximately two weeks after we sent the initial email, we sent a second email telling the decision-maker we no longer needed to speak with them thus releasing them from any obligation to continue trying to reach us. The vast majority of calls from principals are made within the first week of the original email being sent.

4 Empirical Strategy

Our main outcome of interest is whether a decision-maker calls the female parent, the male parent, or neither parent. To test whether our treatments have any effect on the relative proportions of no call, calling the female parent first, or calling the male parent first, we run the following multinomial logit regression:

$$y_i = \beta_j^{l\boldsymbol{M}}(lowMale) + \beta_j^{l\boldsymbol{M}}(highMale) + \beta_j^{l\boldsymbol{F}}(lowFemale) + \beta_j^{h\boldsymbol{F}}(highFemale) + \alpha X_i + \epsilon_i \tag{10}$$

In this regression model, y_i is our outcome variable. In our main analysis, y_i takes one of three values: no call, called female, or called male. We next have treatment indicators for each of non-baseline treatments: lowMale, highMale, lowFemale, and highFemale. We can also include a vector X_i of covariates including which parent the email was sent from (CC'ing the other parent), attributes of the decison-maker and their school. In subsequent analysis we let y_i be a binary variable for ease of

interpreting coefficients.

4.1 Mapping treatment effects to reduced form and structural parameters

When we do not include the vector of covariates X_i, it is straightforward to map the coefficients from the treatment effects regression in Equation (10) to the reduced-form parameters from Equation 4. Further, we can map the treatment effects to the more general reduced-form equation that includes impacts of signals on the beliefs about both parents in Equation (28). This equation is developed in Appendix G.4 and is reproduced in the next paragraph.

We run an unordered logit where $y_i = 0$ if decision-maker i called neither parent; $y_i = 1$ if they called the female parent; and $y_i = 2$ if they called the male parent. Taking calling neither parent as the baseline, we have the following equation for calling the female parent:

$$y_i = \alpha_f + \eta_f^F w_{i,f} + \eta_f^M w_{i,m} + \gamma_f^F w_{i,f} x_{i,f} + \gamma_f^M w_{i,m} x_{i,m} + \varepsilon_{i,f}.$$

We also have the analogous equation for calling the male parent.

Notice that both which parent is called and which parent the message is about matter. 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 adding together the impact of receiving any signal and the impact of the signal value being -1.

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

To build intuition, we show here how to map the treatment effects and reduced-form parameters to the proportions of decision makers in the relevant outcome-signal pairs. For instance,

$$\beta_f^{h\boldsymbol{M}} = \eta_f^M + \gamma_f^M = \ln p_f^{h\boldsymbol{M}} - \ln p_n^{h\boldsymbol{M}} - (\ln p_f^b - \ln p_n^b)$$

That is, the impact of the high signal about the male parent on the calls to female parents is determined as the difference in log proportions of calls to female parents and no parents under the high signal about males versus the baseline log difference of calls to female parents versus male parents.[21]

5 Results

We are balanced on observable variables across our treatments as shown in Appendix Table D3. Although we had intended to send an equal number of emails from fathers and mothers as well as equal number of emails in each of our treatments, these design choices were not attained due to some computing errors.[22] Our main results are based on re-weighted data such that within each of our five messages (Figure 1) there is balance between the number of messages sent from fathers versus mothers. However, our results are quantitatively and qualitatively the same when we randomly exclude observations to achieve balance as shown in Appendix K.

5.1 Gender Inequality With No Signal

Figure 2 presents the proportion of actions taken by decision-makers in our baseline when there is no information about parents' availability. If there was no gender in-

[21] The analogous relationship for the low signal is $\beta_f^{l\boldsymbol{M}} = \eta_f^M - \gamma_f^M = \ln p_f^{l\boldsymbol{M}} - \ln p_n^{l\boldsymbol{M}} - (\ln p_f^b - \ln p_n^b)$.

[22] The issue arose due to the use of the "set seed" command in Stata but was not detected until after our experiment had been fully run. We have no reason to believe that this computing error has introduced any systematic bias into our results.

equality and decision-makers were randomly choosing which parent to call, we would expect the same proportion of calls to male and female parents. In line with Hypothesis 1 outlined in Section 2 we observe that 12% of school principals call mothers first, while only 8% call fathers first. The remaining 79% of decision-makers do not call either parent.[23] The difference in calls to male and female parents is large and statistically significant. Thus we observe a clear gender gap when no signals are given to decision-makers, with mothers being significantly more likely than fathers to be called first.

Another way to see this bias towards calling female parents is in the ratio of female-male calls which is 1.4. This is well above the ratio of 1 which we would expect if decision-makers were randomizing which parent to call and means that mothers are 1.4 times more likely to receive a call than fathers. Conditional on receiving a call back, mothers are called first 59% of the time.

We suspect that the gender gap that we document is a lower bound on the overall gender inequality in external demands for several reasons. First, the type of question in our messages is not a stereotypical male or female question. We would expect external-decision makers to exhibit an even stronger bias towards calling female parents if they needed to call a parent to pick up a sick child, discuss allergies, or help with a bake sale. Second, the school setting itself is universal and as such is not a stereotypical male or female domain. We would expect even more inequality towards mothers if we had sent this same type of message to a doctor's office or dance school because health and dance are stereotypical female domains. In contrast, note that we might expect less of a bias towards mothers if we sent an inquiry about joining a hockey league or to a school but about additional fees, because both are stereotypical male domains. We explore differences by domain in Section 5.4. However, joining an extra-curricular team or paying additional fees (especially at a public school) are not as universal as the experience of being called to pick up a sick child. Furthermore, picking up a sick child is usually an unexpected event that causes a large interruption

[23] A response rate of 20% seems in line with previous work. Recent studies where job applicants submit applications with a phone number and email to an employer find response rates from employers of about 8 to 11% (Agan and Starr, 2018). For work on a similar subject pool of school principals, in line with our expectations, the response rate by phone is lower than the response rate via email observed by others which ranges from 40% to 63% (Bergman and McFarlin Jr, 2018; Ahmed et al., 2020; Oberfield and Incantalupo, 2021). 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, Neal, and Piteo, 2020).

to a person's day, in contrast to less time intensive and more flexible requests about an extra-curricular team or school fees. As such, we believe that the inequality that we document in our setting is a lower-bound on the inequality in external demands on the mother rather than the father when there is no signal about which parent to contact.

5.2 Can Signals Affect Gender Inequality in External Demand for Parents' Time?

5.2.1 Explicit Signals About Availability

Figure 3 shows the proportion of calls made to female and male parents alongside no calls in panel (a) and conditional on a call being made in panel (b). It is clear from Figure 3 that the signals about high and low availability change which parent receives a phone call and can either increase or decrease the baseline bias towards calling female parents.

To rigorously assess the effects of our messages with signals on bias towards calling mothers in comparison to the baseline message Figure 4 visually represents the outcomes from a multinomial logit model like that in Equation 10 (see Table A1 for more details and this same model with and without control variables included). This same multinomial logit model allows us to decompose the mechanisms for the gender inequality into statistical discrimination versus other mechanisms, which we discuss in Section 5.3.

Recall that we randomly vary signals about availability across four messages: HighMale, LowMale, HighFemale, and LowFemaleThere. Two of these messages (HighMale and LowFemale) go against preexisting gender norms by stating that the father has a lot of availability or the mother has limited availability. Figures 3 and 4 show that these messages cause calls to move away from mothers and towards fathers mitigating the gender gap in external demands. The HighMale message reverses the inequality so that mothers are now called 26% of the time while the LowFemale message moves mothers and fathers close to parity with mothers getting 47% of calls and fathers the remaining 53% (Table 1).

In contrast, the remaining two messages, the LowMale and HighFemale, affirm the gender norm that mothers are more available than fathers. We find that they exacerbate the existing inequality by pushing calls towards mothers and away from

fathers. Specifically, stating that the father has low availability results in mothers being called 73% of the time, representing a 24% increase in calls to mothers from the baseline. This change is almost symmetric to the 20% decline in calls to mothers from baseline caused by the LowFemale treatment.

Our results also highlight a striking asymmetry in the effect of our informational interventions. Notably, the HighFemale message stating the mother has high availability results in her being called almost 90% of the time, which is in contrast to father's getting at most about 74% of calls under the HighMale message. Thus, there appears to be a ceiling on how much the father can become the primary point person for external demands, while no such ceiling exists for demands on mothers.

In general, our messages about low availability have smaller effects than those about high availability. We use the effect sizes from Table A1 to identify the exact values of parameters $\bar{r}_f$ and $\bar{r}_m$.

Finally, it is possible that our messages, especially the signals about low availability, might be impacting principals' response rates. We check whether there is any variation in the no call rate across our treatments and find that all of them result in a similar no call rate between 78% to 79% (Table 1 and Figure 4).

5.2.2 Non-verbal Signals: Does Having Fathers Send the Email Decrease Inequality?

In our experiment we randomly vary verbal cues about which parent is more or less available. Our messages have large effects with the HighFemale message resulting in 19% of principals calling the mother versus the HighMale message having only 5% choosing to call the mother; that is a 14 percentage point difference which reverses the gender inequality in favor of men (Table 1). However, there are also non-verbal cues that households can use to signal which parent is the primary point of contact. In our study we randomly assign whether an email comes from the female parent with the male parent is CC'ed or vice-versa. The person sending the email is a non-verbal signal of which parent to contact first.

Pooling across our treatment messages in the Main Variation, we find that the no call rate is similar for both types of senders suggesting that principals are as likely to respond to an email regardless of the identity of the sender (see Tables 3 and Table K10).

However, whether the email is sent by the mother of the father significantly im-

pacts the gender gap in response. Specifically, sending an email from the mother results in her being called 17% and calling the father only 4%, a 13 percentage point difference (similar to what we see between our HighFemale, 19% call mom, and HighMale messages, 5% call mom). In contrast, sending the message from the father results in him being called 14% and the mother only called 7%, a 7 percentage point difference (smaller than the difference between our HighFemale and HighMale messages).

It is clear that the identity of the person sending the email has a large positive effect on who gets the first call. However, that effect is not symmetric for mothers and fathers. Conditional on a call being made, sending the message from the father results in him being called 65% of the time (Figure 5 and Table 3 Panel B Column 1 $MaleNum$), which means external decision makers are still calling the mother one-third of the time even when she didn't send the message. However, when the mother sends the message, 81% of the responding principals call the mother first (Figure 5 and Table 3 Panel A Column 1 $FemaleNum$), resulting in the father being called less than one-fifth of the time. This highlights a ceiling that decision makers have on how much they will let a father be the primary contact for child-related tasks.

Examining the differences across treatment messages in more detail, we see that three of our messages, Baseline/LowMale/HighFemale, result in the mother being called over 95% of the time when she sends the email (bottom three rows of Figure 5). In contrast, none of our messages push the father to be called more than 95% of the time when he sends the email. This underscores a striking asymmetry in the effects of informational interventions on the gender gap in external demand for parental involvement and suggests that external decision makers have a ceiling on how much they will contact the father, while no such ceiling exists for mothers.

Last, something striking about Figure 5 is that almost none of our email-treatment pairs result in a 50-50 split between calls to mothers and fathers. This is in spite of many households reporting they would like closer to equal splits in parenting responsibilities. This may be because most schools, and other child related activities, only allow two-parent households to denote a "Contact 1" and "Contact 2", essentially pushing the household towards a corner solution of always call mom or always call dad. This is likely an artifact of traditional gender norms where one-parent specializes in household production, while the other in outside the home production. The database systems that schools use push households towards a corner solution, when

many households would prefer an interior solution.

5.3 What Drives Gender Inequality?

Our theoretical model described in section 2 allows us to investigate whether the gender inequality we observe in the *baseline* message is driven by decision-maker's beliefs about parents' responsiveness or other deterrents. Intuitively, in the US, mothers are more likely to be stay-at-home parents than fathers.[24] This general statistical information could lead decision-makers to believe that female parents on average will be more responsive, and as such will bias decision-makers toward making more external demands of women. In Appendix F we show that these types of decision-makers indeed report that they prefer to contact female parents because they believe mothers are more responsive, but also because they believe mothers may be the primary contact about child related topics which we address in section 5.3.1. Furthermore, in our own survey we find that female parents self-report being the first to respond 82% of the time, while male parents are the first to respond to the school only 42% of the time, indicating that decision-makers correctly anticipate that mothers are more responsive. In section 6.2 we discuss whether contacting the female parent first is in fact efficient under various definitions of efficiency.

Beyond responsiveness, there may be other deterrents affecting decision makers' choice to call a parent of certain type. For example, they may prefer talking to mothers because they are more pleasant, or prefer talking to fathers because they are better able to make decisions for the whole household in a patriarchal society. Alternatively, they may decide which parent to call based on the prevailing gender norms. There may also be other belief-based factors, unrelated to responsiveness. For example, in our specific setting, principals may believe that mothers are easier to convince to enroll in their school which may explain why they are more likely to call mothers than fathers. Finally, institutional or systemic discrimination may also lead to the gender gaps that we observe. While we cannot disentangle the role of each possible factor in our experiment, we can shed light on the relative role of beliefs about responsiveness vis-a-vis other deterrents.

We find that our parameter estimate for responsiveness of female parents is $\bar{r}_f = -0.34$ which is less than the analogous parameter for male parents $\bar{r}_m = -0.25$,

[24] `https://www.census.gov/content/dam/Census/library/visualizations/time-series/demo/families-and-households/shp-1b.pdf`

although the difference is not statistically significant ($Prob > chi2 = 0.60$). These parameters are constructed from the results of the multinomial logit reported in Table A2. This hints that although some of the gender inequality we observed from our baseline messages is driven by decision-makers' beliefs about responsiveness as outlined in Hypothesis 2, beliefs about responsiveness do not seem to be the main driver of the gender inequality in external demands for parents' time.

Next, we test if the gender inequality we document can be explained by other deterrents, as discussed by Hypothesis 3. We find that our parameter estimates for the residual term for male parents is greater than that for female parents, that is $\bar{\delta}_m - \bar{\delta}_f = 0.48$ ($Prob > chi2 = 0.0008$). This is direct evidence that some of the gender inequality in demand for parents' involvement is driven by factors unrelated to beliefs about responsiveness. We investigate these factors below.

5.3.1 Beliefs About Parental Involvement and Expertise

It is likely that in deciding which parent to call, decision-makers want to get not only a quick but also useful response. Indeed in our own survey, educators reported that they wanted to call mothers most often both because mothers were more responsive and because mothers were more likely to be the primary person making decisions about a wide range of child-related topics (e.g., sick child, child's allergies, school related payments, volunteering at a book fair or at career day).

To better understand if our findings are partially driven by beliefs that child-related choices are primarily made by mothers, we added the following sentence to all our messages "This is the type of decision we both want to be involved in equally." We sent out an additional 30,320 emails with this additional sentence which we call the "Equal Involvement" variation.

If beliefs that mothers primarily make child related decisions are driving some of the inequality, then we would expect fewer calls to mothers with the addition of this sentence. As detailed in Table 2 we find that mothers receive 12.0% of the calls and fathers 8.0% of the calls in the "Equal Decision" variation (see Appendix J for details by message variations). Conditional on a call being made, mothers get 60.0% of calls in the "Equal Involvement" variation, which is almost identical to our findings in the Main variation of 12.2% of calls to mothers, 8.8% of calls to fathers, and conditional on a call 58.1% of calls to the mother. Some of these differences (12.2% vs. 12.0% and 8.8% vs. 8.0% and 58.1% vs. 60.0%) are statistically significant but not economically

significant. Overall, we see the ratio of calls to mothers versus fathers slightly rising in the "Equal Involvement" variation relative to the Main variation, which makes it all the less likely that our findings are driven by beliefs that mothers primarily make child related choices.

5.3.2 Beliefs About Stay-at-home Mothers

In the U.S., mothers are significantly more likely to be stay-at-home parents than fathers.[25] To better understand if our findings are partially driven by beliefs about stay-at-home parents being more likely to be female we added the following sentence to all our messages "We both work full-time." This is meant to shut down belief-based mechanism that the mother is a stay-at-home parent. We call this the "Full-time" variation and we sent this to an additional 9,472 principals (see Appendix J for details by message variations).

We would expect fewer calls to mothers in our "Full-time" variation if beliefs that mothers were more likely to be a stay-at-home parent are driving the gender inequality. We do not find evidence of this as shown in Table 2. The rates of calls to mothers and fathers are quite similar in the Full Time variation and the Main variation. In the Full Time variation mothers receive 11.9% of the calls and fathers receive 7.3% of the calls, which is almost identical to the Main variation. Conditional on a call being made, the mother is called 62.0% of the time. In fact the ratio of calls to mothers versus fathers rises very slightly from 58% in the Main variation when we include information that shuts down the idea that the mother is a stay-at-home parent.

5.3.3 Gender Norms

Another mechanism that could explain the gender gap in external demand for parental involvement that we document in our experiment is a strong gender norm governing interactions between decision makers and parents. For example, in our own survey, school educators report being more likely to call mothers because "mothers are more caring and nurturing," "more excited to participate," and "more polite." They also report having a better relationship with mothers.

[25] In 2016, about 27% of households had a stay-at-home mother while only 7% of households had a stay at home father. See `https://www.pewresearch.org/fact-tank/2018/09/24/stay-at-home-moms-and-dads-account-for-about-one-in-five-u-s-parents/`

As prior studies have documented, despite women's considerable advances in education and labor market outcomes in recent years, social norms about gender identity have been quite persistent and still impact a wide range of economic and social outcomes from labor force participation and earnings to marriage formation, fertility, and division of home production (Bertrand, Kamenica, and Pan, 2015a; Kerwin, Guryan, and Pan, 2022). We investigate whether gender norms can partly explain principals' behavior by linking our school data to a state-level sexism index constructed by Kerwin et al. (2022) from the General Social Survey (GSS) data measuring gender attitudes.

We find suggestive evidence that gender norms are an important mechanism in explaining the gender inequality in external demand for parental involvement that we document in our setting. While we find only marginally more separation between $\bar{r}_f$ and $\bar{r}_m$ in areas with more sexist beliefs, we lack precision in our state-level sexism measures. These findings, however, are suggestive that gender attitudes are related to decision makers' beliefs about parents' relative availability in our setting.

Furthermore, when we compare private schools that identify as religious with non-religious private and public schools, we find larger inequality in calls to mothers vs. fathers. In particular, baseline unconditional call-back rates for religious schools are 23% to moms and 7% to dads, versus 11-14% and 7-8% for public+private (non religious) schools, respectively.

5.4 Is Inequality Lower in a More Male-Domained Task?

It is possible that both male and female parents are fielding similar numbers of external requests, but that certain types of requests are associated with the female or male domain. Our own survey (Section F) found that within the school setting educators stated they most heavily favored calling the mother for a child being sick, volunteering at a book fair and when dealing with allergies. While the educators still favored the mother, they did so to a lesser degree for requests to volunteer for a career day and to discuss school payments.

To test if fathers are contacted more often in more male-stereotyped domains we fielded an additional variation of our email messages which states "We are searching for schools for our child and are especially interested in discussing school fees and other expenses." In this variation, we observe fewer calls to mothers with only 9.7%

of principals calling the female parent when the sentence about fees is added (vs. 12.2% in the Main variation $p = 0.00$). However, we also see fewer calls to fathers with only 7.0% of principals calling fathers in response to the emails about fees (vs. 8.8% $p = 0.00$). The actual rate of calling mothers versus fathers conditional on a call being made is not statistically significantly different from the Main variation at 58.0% (vs. 58.1%). Thus, even in a stereotypically male domain within the school setting we do not see a shifting of the calls from mothers to fathers.

5.5 Heterogeneous Treatment Effects

A better understanding of which types of decision-makers and schools are most susceptible to gender inequality will both increase our understanding of mechanisms as well as our ability to target interventions. This section is pending the full matching of our call data with the schools' data as detailed in Section E.

6 Discussion and Conclusion

6.1 Generalizability of Results

Our theoretical model provides intuition for the underlying drivers behind the striking gender inequality in demand for parental involvement that we document in our field experiment. The framework can be adapted to derive results in other settings, where the forces at work could differ.

Setting: As a first step towards understanding gender inequality in external demands, we conduct a field experiment in the K-12 the school setting. This setting is broadly applicable to the general population since 40% of households in the US have school-aged children. However, it is possible that our results could differ by the setting with some settings showing a similar skew towards calling mothers (e.g. doctors, dentists, daycare programs, extra-curricular activities, summer camps) and others likely showing a skew towards fathers (e.g. household retirement planning, Boy Scouts). Furthermore, because we are interested in gender inequality, we only study two parent households with one male and one female parent, while other household arrangements exist and are important to study. We discuss these in Appendix B.

Lower Bound: We believe that the inequality we document is likely a lower bound on the total inequality that women face in external demands on their time versus men. It is likely that women experience more interruptions regarding not only the needs of their children, but also the needs of any adult family members who require care taking.[26] Also, researchers are increasingly documenting that women shoulder a disproportionately large share of the cognitive load associated with managing a household (Daminger, 2019). Activities such as coordinating childcare, thinking about and anticipating future household needs, and other forms of invisible mental labor tend to be highly gendered and impose substantial disruptions to women's paid work.

Using the language of List (2020), this study represents a "first wave" study in which we focus on establishing causality and illuminating mechanisms with the help of a theoretical model. Although our evidence comes from a particular setting (schools) and a specific decision-maker (school principal), our conceptual framework and research design can be adapted to other settings and to adjacent research questions.

6.2 Efficiency

There are multiple parties involved in the interaction that we have investigated: the parents, the external decision-maker (in our case the school), the child, and if the parents are employed, their employers. With multiple parties involved and many trade-offs to consider, it is not readily apparent what is the most efficient allocations of calls between mothers and fathers. We discuss this below.

External Decision-Makers: Decision-makers may have multiple competing objectives. In our model (Section 2), the decision-maker in the short run is maximizing the likelihood of a useful response. However, in the long run, an entity (school, church, extra-curricular program, doctor) may find it desirable to have a more diverse set of parents involved (e.g. not skewed towards mothers), and they may also prefer to have more parents (e.g. both parents vs. one parent) involved (Clark, Lotto, and McCarthy, 1980). A less myopic decision-maker may want to call fathers even if they believe the father is less likely to respond or may provide a less useful response. We believe work on these trade offs is an important area for future research.

[26] `https://www.aarp.org/ppi/info-2020/caregiving-in-the-united-states.html`

Parents: Survey evidence indicates that mothers on average wish they were contacted less often about child-related needs than they currently are (see Appendix F.1). The existing skew towards 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 likely would reduce parents' stress levels; such reductions in stress are associated with better parenting (Conger, Conger, and Martin, 2010).

In our experimental data, even when the email comes from the father and that father signals they have high availability, 12 percent of calls are still directed to mothers (Table 3). 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. For all these reasons, the current inequality in demands for parental involvement appears to be inefficient for parents.

Child: The skew toward mothers being called more may be welfare-harming for children as there is increasing evidence that children benefit from having both fathers and mothers involved (Pleck, 2007).

Parents' Employers and Economic Efficiency: Parents' employers would like to minimize interruptions to their employees' work day. If the school is going to contact a parent, each employer would prefer that the school contact 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 the collective of mother's and father's workplaces (and the overall economy) for the parent who has signalled 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 that they do not fully integrate them, as 26% of calls still go to mothers even when the father states he is highly available (Table 1).

Further investigation of the trade offs each party faces, and how a social planner

might weigh the needs of the various parties is an important next step in this research agenda.

6.3 Conclusion

In this paper, we investigate a novel gender inequality in external demands for parental involvement. We develop a theoretical model which motivates the design of a large-scale field experiment in a K-12 school setting. In this experiment, we send emails to over 80,000 U.S. school principals with a general inquiry about the school and a request to call one of the parents back. We randomly vary signals about parents' availability as well as which parent sends the email.

We document a striking gender inequality in responses. Conditional on receiving a callback, mothers are called first 40% more than fathers. To our knowledge, this provides the first empirical evidence of a significant gender inequality in external demands for parental time. We show that signaling father's availability mitigates this inequality and causes mothers to be called less than half the time. However, we observe a striking asymmetry in the effect of our informational interventions. Specifically, even when fathers strongly signal their availability, mothers are still called 26% of the time. In contrast, signals that reinforce stereotypes about mothers being more available cause mothers to receive 90% of the calls. Notably, even when the email comes from the father *and* the father signals his availability, 12% of the calls are still directed to mothers. In contrast, fathers receive only 3% of the calls when mothers are the primary senders and signal that they are available. This underscores a ceiling on the degree to which informational signals can mitigate gender inequality in external demand for parental involvement.

Our theoretical model allows us to disentangle the mechanisms underlying any differential demand for parental time into beliefs about responsiveness versus other deterrents. We measure the impact of beliefs about responsiveness by randomizing the signals we send to decision makers about the availability and/or involvement of a specific parent while the other factors are measured as a residual term in our model. We find that decision-makers hold similar beliefs about the responsiveness of male and female parents in our setting. In contrast, we find that the inequality that we document is driven in part by differences in the residual. We test several potential deterrents, including beliefs about mothers being more likely to be stay at home

parents, being the primary decision-maker on child-related choices, and the role of gender norms. We find evidence that gender norms are in part responsible for the gender gap in external demand for parental involvement.

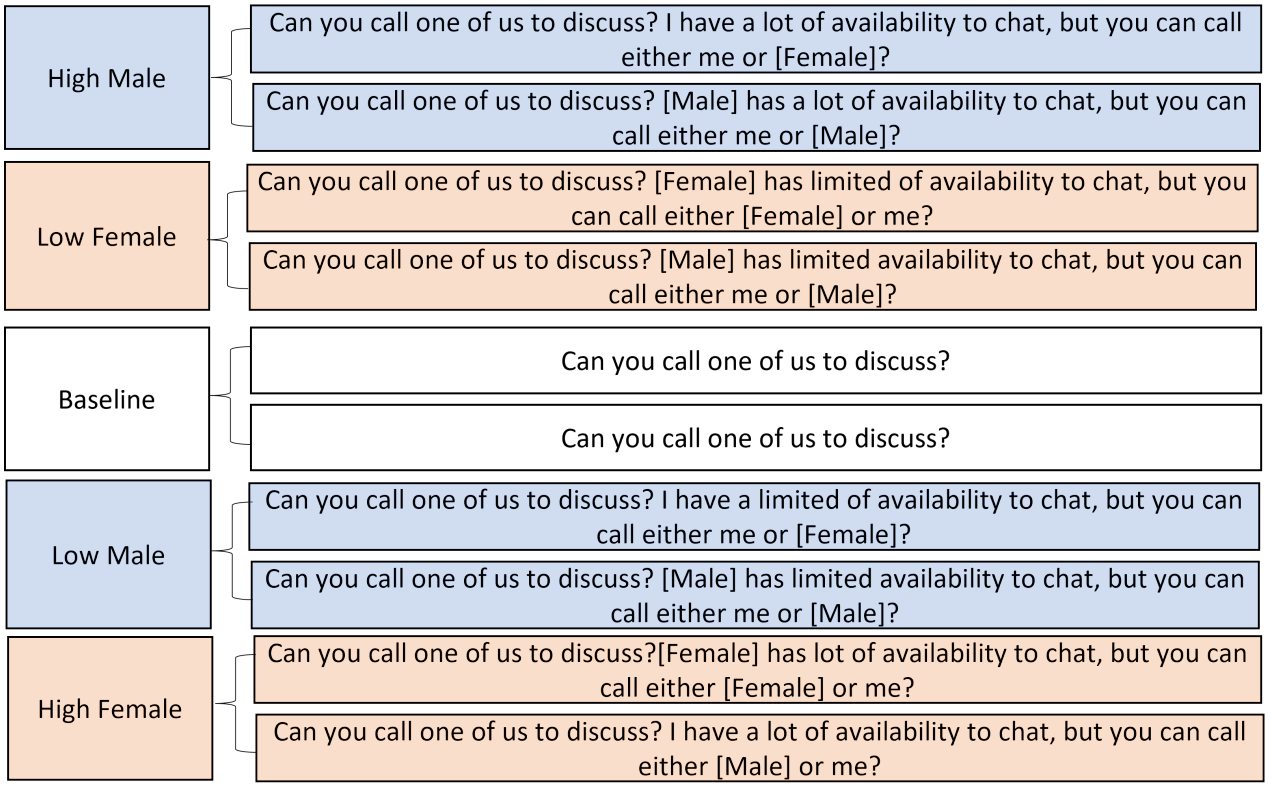
We believe that the patterns we document represent a lower bound on the overall gender inequality in demands for parental involvement. The school setting is only one of many domains where gender differences in external demands on parents' time lead to disproportionate workday interruptions for mothers. While it is possible that fathers receive more requests in certain (male-stereotyped) domains, we do not observe this to be the case even in the most male-domained task in our experiment (asking about school payments).

The gender gap that we document can have detrimental and persistent effects on women's career trajectories. More frequent workday interruptions for women versus men have been linked to a wide range of important economics outcomes, including occupational choice, human capital accumulation, and promotions. Furthermore, if women are disproportionately shouldering child-related and household tasks, they incur substantial personal costs, including to physical and mental health. Investigating the source of these inequalities and documenting that they are in part driven by external demands informs policies aimed at mitigating the gaps. As our findings indicate, both households' and external decision makers' actions can affect the size of the inequality. Parents signaling fathers' availability and schools engaging in more equitable involvement of both parents are needed to mitigate the gap.

7 Tables and Figures

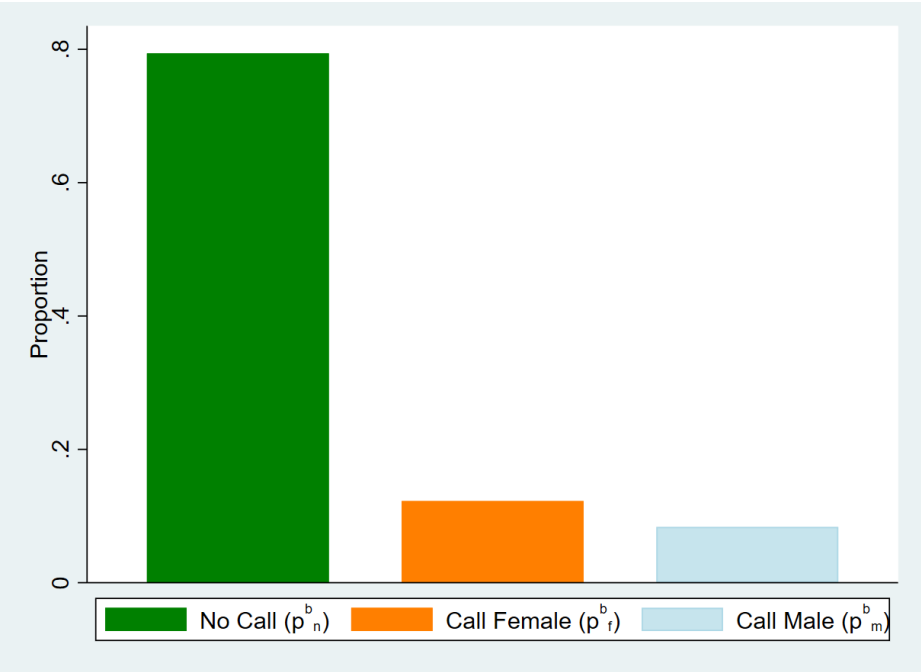
7.1 Figures

Figure 1: Field Experiment Variation In Messages



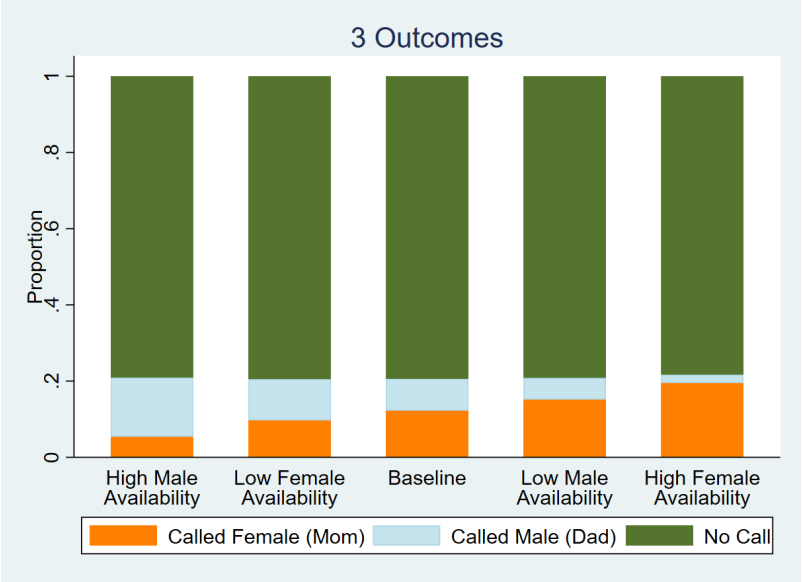
Notes: In this figure we show pertinent portion of variation in the messages we sent to schools. The parent who sent the email always had their phone number listed first. The full text of example email messages is available in Appendix Section I

Figure 2: Outcomes in Baseline

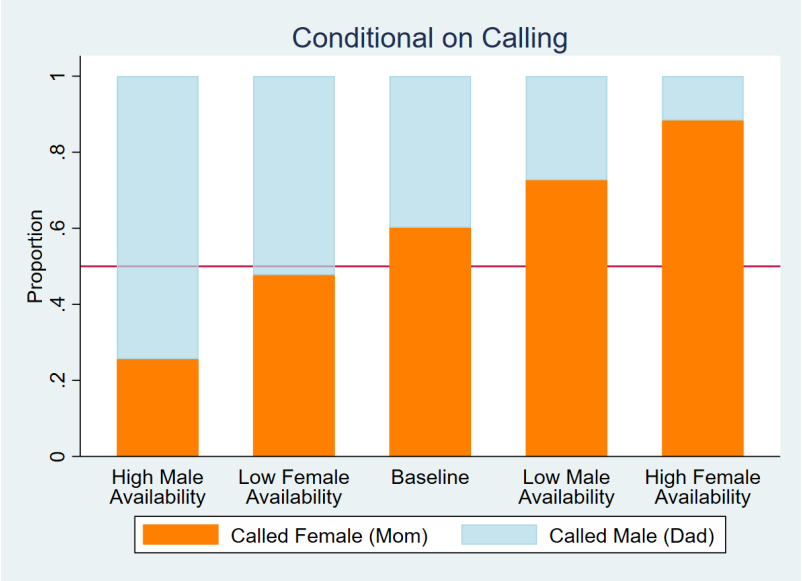


Notes: In this figure we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) when the decision-maker is sent the baseline message with no signal about which parent has high or low availability in our Main Variation ($N = 30,471$). If 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 there is 50% of emails from a female parent and 50% from a male parent.

Figure 3: Outcomes By Treatment



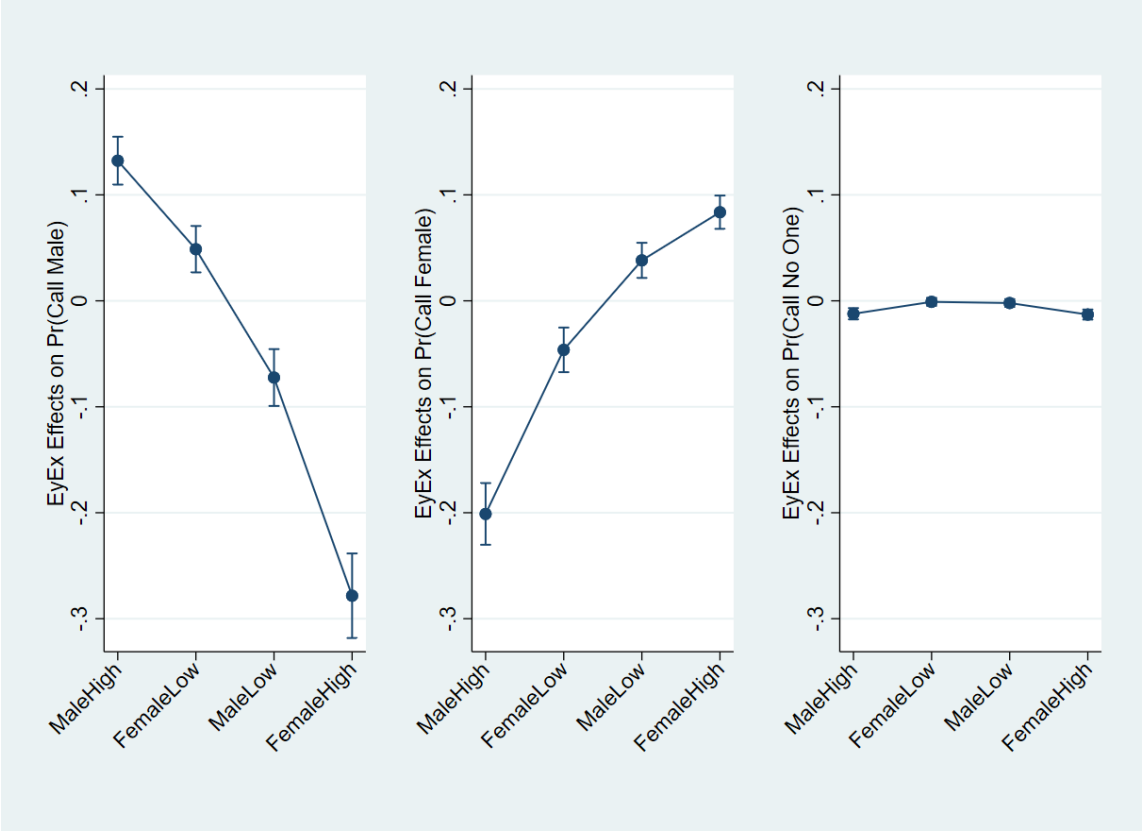
(a) All Outcomes



(b) Outcomes Conditional On Calling

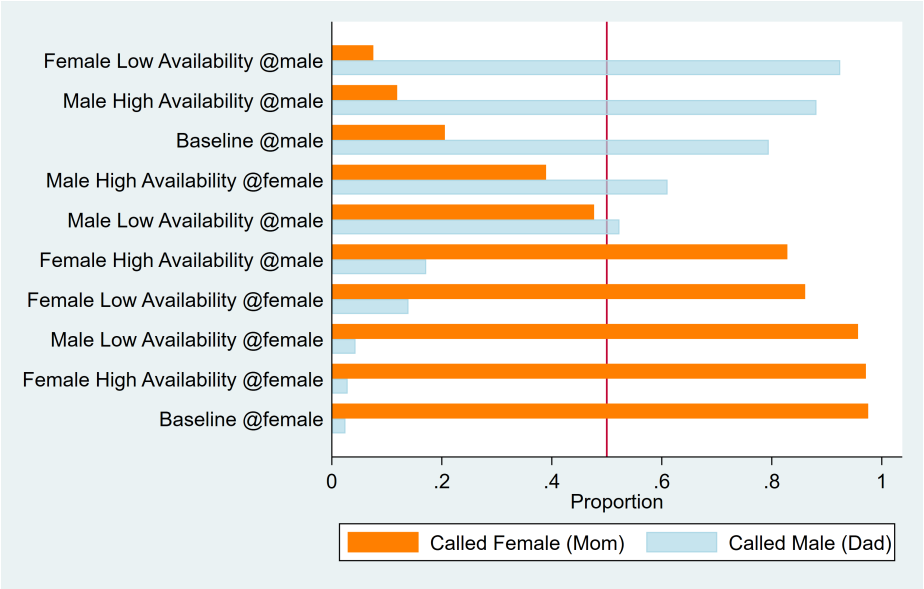
Notes: In this figure we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision-maker in our Main Variation. Panel (a) represents three outcomes from $30,471$ decision-makers, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 7,778$). 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 there is 50% of emails from a female parent and 50% from a male parent.

Figure 4: Effects By Treatment



Notes: In this figure we show the results from a mutltinomial logit model using a model like 10 which is detailed fully in Table A1. This figure shows the marginal effects elasticities. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.

Figure 5: Outcomes in Main By Treatment and Email Sender



Notes: In this figure we show the proportion of decision-makers choosing to call the female parent (mom) or the male parent (dad) conditional on a call being made by Treatment in our Main Variation ($N = 30,471$) and whether the sender of the email was the female or male parent (and CCing the other parent). The ordering from top to bottom is sorted by "proportion called female-proportion called male". A red vertical line is shown at the 50-50 equal split of calls between mothers and fathers, which is almost never where the actual observed data lays. Details of "No Call" are shown in Table 3

7.2 Tables

Table 1: **Summary Statistics By Treatment in Main Variation**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
FemaleNum0	0.05	0.10	0.12	0.15	0.19
MaleNum0	0.16	0.11	0.08	0.06	0.02
NoCall	0.79	0.79	0.79	0.79	0.78
FemaleNum	0.26	0.47	0.59	0.73	0.90
MaleNum	0.74	0.53	0.41	0.27	0.10
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	7075	5931	5612	5700	6153

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.

Table 2: **Summary Statistics By Variation (All Treatments Combined)**

	(1) Main	(2) Equal Decision	(3) Full Time	(4) Payments
FemaleNum0	0.122	0.120	0.119	0.097
MaleNum0	0.088	0.080	0.073	0.070
NoCall	0.791	0.800	0.809	0.833
FemaleNum	0.581	0.600	0.620	0.580
MaleNum	0.419	0.400	0.380	0.420
FemaleEmail	0.500	0.500	0.500	0.500
Observations	30471	30320	9472	9808

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.

Table 3: **Summary Statistics By Primary Email Sender**

Panel A: Email Sent By Mother (CCing Father)

	(1) All Messages	(2) High Male (Hm)	(3) Low Female (Lf)	(4) Baseline (b)	(5) Low Male (Lm)	(6) High Female (Hf)
FemaleNum0	0.17	0.08	0.18	0.20	0.21	0.21
MaleNum0	0.04	0.13	0.03	0.01	0.01	0.01
NoCall	0.79	0.78	0.79	0.79	0.78	0.79
FemaleNum	0.81	0.39	0.86	0.98	0.96	0.97
MaleNum	0.19	0.61	0.14	0.02	0.04	0.03
FemaleEmail	1.00	1.00	1.00	1.00	1.00	1.00
Observations	15560	3712	2726	3108	2895	3119

Panel B: Email Sent By Father (CCing Mother)

	(1) All Messages	(2) High Male (Hm)	(3) Low Female (Lf)	(4) Baseline (b)	(5) Low Male (Lm)	(6) High Female (Hf)
FemaleNum0	0.07	0.02	0.02	0.04	0.10	0.18
MaleNum0	0.14	0.18	0.19	0.16	0.11	0.04
NoCall	0.79	0.80	0.80	0.80	0.80	0.78
FemaleNum	0.35	0.12	0.08	0.21	0.48	0.83
MaleNum	0.65	0.88	0.92	0.79	0.52	0.17
FemaleEmail	0.00	0.00	0.00	0.00	0.00	0.00
Observations	14911	3363	3205	2504	2805	3034

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother. Observations do not have to be weighted in this table by whether the email sender is the mother or father in this table because the panels only show responses to emails from mother (top panel) or father (bottom panel).

References

Adams-Prassl, A. (2023). The gender wage gap on an online labour market: the cost of interruptions. *Review of Economics and Statistics*.

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

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.

Aguiar, M. and E. Hurst (2007). Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics*.

Ahmed, A., M. Hammarstedt, and K. Karlsson (2020). Do swedish schools discriminate against children with disabilities?

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *Ilr Review 30*(2), 175–187.

Albanese, A., A. Nieto, and K. Tatsiramos (2022). Job location decisions and the effect of children on the employment gender gap. *CESifo Paper 9792*.

Amuedo-Dorantes, C., M. Marcén, M. Morales, and A. Sevilla (2020, October). COVID-19 School Closures and Parental Labor Supply in the United States. IZA Discussion Papers 13827, Institute of Labor Economics (IZA).

Anderson, D. J., M. Binder, and K. Krause (2002). The motherhood wage penalty: Which mothers pay it and why? *American economic review 92*(2), 354–358.

Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of labor economics 34*(3), 545–579.

APA (2021). American psychological association.

Arrow, K. (1973). The theory of discrimination,. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*, pp. 3–3. Princeton: Princeton University Press.

Bailey, M. J., T. S. Byker, E. Patel, and S. Ramnath (2019). The long-term effects of california's 2004 paid family leave act on women's careers: Evidence from us tax data.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago press.

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

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. *Handbook of economic field experiments 1*, 309–393.

Bertrand, M., E. Kamenica, and J. Pan (2015a, 01). Gender Identity and Relative Income within Households *. *The Quarterly Journal of Economics 130*(2), 571–614.

Bertrand, M., E. Kamenica, and J. Pan (2015b). Gender identity and relative income within households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bertrand, M. and S. Mullainathan (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review 94*(4), 991–1013.

Bianchi, S. M., J. P. Robinson, and M. A. Milke (2006). *The changing rhythms of American family life*. Russell Sage Foundation.

Bianchi, S. M., L. C. Sayer, M. A. Milkie, and J. P. Robinson (2012). Housework: Who did, does or will do it, and how much does it matter? *Social forces 91*(1), 55–63.

BLS (2021). Bureau of labor statistics.

Bohren, J. A., P. Hull, and A. Imas (2022). Systemic discrimination: Theory and measurement. Technical report, National Bureau of Economic Research.

Bohren, J. A., A. Imas, and M. Rosenberg (2019). The dynamics of discrimination: Theory and evidence. *American economic review 109*(10), 3395–3436.

Boye, K. (2015). Can you stay home today? parents' occupations, relative resources and division of care leave for sick children. *Acta Sociologica 58*(4), 357–370.

Brysbaert, M. (2019). How many words do we read per minute? a review and meta-analysis of reading rate. *Journal of memory and language 109*, 104047.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017, November). 'acting wife': Marriage market incentives and labor market investments. *American Economic Review 107*(11), 3288–3319.

Bygren, M., A. Erlandsson, and M. Gähler (2017). Do employers prefer fathers? evidence from a field experiment testing the gender by parenthood interaction effect on callbacks to job applications. *European Sociological Review 33*(3), 337–348.

Charmes, J. (2019). The unpaid care work and the labour market. an analysis of time use data based on the latest world compilation of time-use surveys. *International Labour Office–Geneva: ILO*.

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.

Correll, S. J., S. Benard, and I. Paik (2007). Getting a job: Is there a motherhood penalty? *American journal of sociology 112*(5), 1297–1338.

Cortes, P. and J. Pan (2016). Prevalence of long hours and skilled women's occupational choices.

Cortes, P. and J. Pan (2021). Children and the remaining gender gaps in the labor market. *Working Paper*.

Couch, K. A., R. W. Fairlie, and H. Xu (2022). The evolving impacts of the covid-19 pandemic on gender inequality in the us labor market: The covid motherhood penalty. *Economic Inquiry 60*(2), 485–507.

Council, N. R. (2004). Measuring racial discrimination.

Craig, L. and K. Mullan (2011). How mothers and fathers share childcare: A cross-national time-use comparison. *American sociological review 76*(6), 834–861.

Cubas, G., C. Juhn, and P. Silos (2021). Work-care balance over the day and the gender wage gap. In *AEA Papers and Proceedings*, Volume 111, pp. 149–53.

Cubas, G., C. Juhn, and P. Silos (2022). Coordinated work schedules and the gender wage gap. *Economic Journal*.

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.

Daminger, A. (2019). The cognitive dimension of household labor. *American Sociological Review 84*(4), 609–633.

Duchini, E. and C. Van Effenterre (2022). School schedule and the gender pay gap. *Journal of Human Resources*.

Eil, D. and J. M. Rao (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics 3*(2), 114–38.

Erosa, A., L. Fuster, G. Kambourov, and R. Rogerson (2022). Hours, occupations, and gender differences in labor market outcomes. *American Economic Journal: Macroeconomics 14*(3), 543–90.

Flabbi, L. and A. Moro (2012). The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model. *Journal of Econometrics 168*(1), 81–95.

Garcia, K. S. D. and B. W. Cowan (2022). The impact of school and childcare closures on labor market outcomes during the covid-19 pandemic. Technical report, National Bureau of Economic Research.

Gicheva, D. (2013). Working long hours and early career outcomes in the high-end labor market. *Journal of Labor Economics 31*(4), 785–824.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Goldin, C. and L. F. Katz (2011). The cost of workplace flexibility for high-powered professionals. *The Annals of the American Academy of Political and Social Science 638*(1), 45–67.

Hansen, B., J. Sabia, and J. Schaller (2022). Schools, job flexibility, and married women's labor supply.

He, H., S. X. Li, and Y. Han. Labor market discrimination against family responsibilities: A correspondence study with policy change in china.

Heggeness, M. L. (2020). Estimating the immediate impact of the covid-19 shock on parental attachment to the labor market and the double bind of mothers. *Review of Economics of the Household 18*(4), 1053–1078.

Islam, A., D. Pakrashi, L. C. Wang, and Y. Zenou (2018). Determining the extent of statistical discrimination: Evidence from a field experiment in india. Available at SSRN: https://ssrn.com/abstract=3185899.

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*.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kleven, H., C. Landais, and J. E. Søgaard (2021). Does biology drive child penalties? evidence from biological and adoptive families. *American Economic Review: Insights 3*(2), 183–98.

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.

Knowles, J., N. Persico, and P. Todd (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy 109*(1), 203–229.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018). The mommy effect: Do women anticipate the employment effects of motherhood? Technical report, National Bureau of Economic Research.

Laouénan, M. and R. Rathelot (2022). Can information reduce ethnic discrimination? evidence from airbnb. *American Economic Journal: Applied Economics 14*(1), 107–32.

Lee, Y.-S. and L. J. Waite (2005). Husbands' and wives' time spent on housework: A comparison of measures. *Journal of marriage and family 67*(2), 328–336.

List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics 119*(1), 49–89.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, NBER.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*(12), 3722–59.

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

Montes, J., C. Smith, and I. Leigh (2021). Caregiving for children and parental labor force participation during the pandemic.

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.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources 47*(4), 1128–1157.

Nunley, J. M., A. Pugh, N. Romero, and R. A. Seals (2015). Racial discrimination in the labor market for recent college graduates: Evidence from a field experiment. *The BE Journal of Economic Analysis & Policy 15*(3), 1093–1125.

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.

Offer, S. (2014). The costs of thinking about work and family: Mental labor, work–family spillover, and gender inequality among parents in dual-earner families. In *Sociological Forum*, Volume 29, pp. 916–936. Wiley Online Library.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? a field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy 3*(4), 148–71.

Pertold-Gebicka, B., F. Pertold, and N. Datta Gupta (2016). Employment adjustments around childbirth.

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the french financial sector. *Labour Economics 14*(3), 371–391.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review 62*(4), 659–661.

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

Powell, W. W. and P. J. DiMaggio (2012). *The new institutionalism in organizational analysis*. University of Chicago press.

Price, B. M. and M. Wasserman (2022). The summer drop in female employment.

Russell, L. and C. Sun (2020). The effect of mandatory child care center closures on women's labor market outcomes during the covid-19 pandemic.

Sarsons, H. (2017). Interpreting signals in the labor market: evidence from medical referrals.

Scarborough, W. J., R. Sin, and B. Risman (2019). Attitudes and the stalled gender revolution: Egalitarianism, traditionalism, and ambivalence from 1977 through 2016. *Gender & Society 33*(2), 173–200.

Schoonbroodt, A. (2018). Parental child care during and outside of typical work hours. *Review of Economics of the Household 16*(2), 453–476.

Scott, W. R. (2013). *Institutions and organizations: Ideas, interests, and identities*. Sage publications.

Sevilla, A. and S. Smith (2020). Baby steps: The gender division of childcare during the covid-19 pandemic. *Oxford Review of Economic Policy 36*(Supplement_1), S169–S186.

Small, M. L. and D. Pager (2020). Sociological perspectives on racial discrimination. *Journal of Economic Perspectives 34*(2), 49–67.

Smith, T. W., M. Davern, J. Freese, and S. L. Morgan.

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

Wasserman, M. (2022). Hours constraints, occupational choice, and gender: Evidence from medical residents. *Review of Economic Studies*.

Wikle, J. and C. Cullen (2021). The developmental course of parental time investments in children from infancy to late adolescence. *Working Paper*.

Wiswall, M. and B. Zafar (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics 133*(1), 457–507.

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.

Appendix: For Online Publication Only

A Appendix Tables

Table A1: **Multinomial Logit Models of Effect of Treatments on No Call, Call Male or Call Female**

	(1) outcome	(2) outcome	(3) outcome	(4) outcome	(5) outcome	(6) outcome
No_Call						
High Male (Hm)	-0.62***	-0.66***	0.81***	0.85***	0.00	0.00
	(0.06)	(0.06)	(0.07)	(0.07)	(.)	(.)
Low Female (Lf)	-0.26***	-0.27***	0.23***	0.24***	0.00	0.00
	(0.07)	(0.07)	(0.06)	(0.06)	(.)	(.)
Low Male (Lm)	0.38***	0.39***	-0.22***	-0.23***	0.00	0.00
	(0.08)	(0.08)	(0.05)	(0.05)	(.)	(.)
High Female (Hf)	1.31***	1.35***	-0.48***	-0.51***	0.00	0.00
	(0.10)	(0.10)	(0.05)	(0.05)	(.)	(.)
Female_Call						
High Male (Hm)	-1.44***	-1.51***	0.00	0.00	-0.81***	-0.85***
	(0.08)	(0.09)	(.)	(.)	(0.07)	(0.07)
Low Female (Lf)	-0.49***	-0.51***	0.00	0.00	-0.23***	-0.24***
	(0.08)	(0.08)	(.)	(.)	(0.06)	(0.06)
Low Male (Lm)	0.59***	0.62***	0.00	0.00	0.22***	0.23***
	(0.09)	(0.09)	(.)	(.)	(0.05)	(0.05)
High Female (Hf)	1.79***	1.86***	0.00	0.00	0.48***	0.51***
	(0.11)	(0.11)	(.)	(.)	(0.05)	(0.05)
Male_Call						
High Male (Hm)	0.00	0.00	1.44***	1.51***	0.62***	0.66***
	(.)	(.)	(0.08)	(0.09)	(0.06)	(0.06)
Low Female (Lf)	0.00	0.00	0.49***	0.51***	0.26***	0.27***
	(.)	(.)	(0.08)	(0.08)	(0.07)	(0.07)
Low Male (Lm)	0.00	0.00	-0.59***	-0.62***	-0.38***	-0.39***
	(.)	(.)	(0.09)	(0.09)	(0.08)	(0.08)
High Female (Hf)	0.00	0.00	-1.79***	-1.86***	-1.31***	-1.35***
	(.)	(.)	(0.11)	(0.11)	(0.10)	(0.10)
Control Variables		Yes		Yes		Yes
R^2						
Observations	30471	30471	30471	30471	30471	30471

Notes: This table presents the results of a multinomial logit model using a model like the one in Equation 10. 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 in columns (1) and (2), female call in columns (3) and (4), and male call in columns (5) and (6). The results from the three base cases are analogous and all three are presented to make specific comparisons more simple. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.The outcomes with no controls from this table are represented visually in Figure 4.

Table A2: **Multinomial Logit Models For Theory Model**

	(1) Main	(2) Equal Decision	(3) Full Time	(4) Payments
No_Call				
any_msg_M	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
x_M	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
any_msg_F	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
x_F	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
Female_Call				
any_msg_M	-0.30***	-0.19***	-0.28**	-0.37***
	(0.05)	(0.05)	(0.10)	(0.10)
x_M	-0.51***	-0.36***	-0.43***	-0.59***
	(0.03)	(0.03)	(0.07)	(0.07)
any_msg_F	0.12*	0.31***	0.36***	0.04
	(0.05)	(0.05)	(0.09)	(0.10)
x_F	0.36***	0.36***	0.27***	0.35***
	(0.03)	(0.03)	(0.05)	(0.06)
Male_Call				
any_msg_M	0.12*	0.20***	0.28**	0.29*
	(0.06)	(0.06)	(0.10)	(0.12)
x_M	0.50***	0.41***	0.41***	0.55***
	(0.03)	(0.03)	(0.06)	(0.06)
any_msg_F	-0.53***	-0.34***	-0.55***	-0.40**
	(0.07)	(0.06)	(0.12)	(0.15)
x_F	-0.78***	-0.49***	-0.74***	-0.80***
	(0.05)	(0.04)	(0.09)	(0.11)
Control Variables				
R^2				
Observations	30471	30320	9472	9808

Notes: This table presents the results of a multinomial logit model using a model like the one in Equation 10. 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. The right hand side variables are discussed in Section 2. Observations are weighted so that there is 50% of emails from a female parent and 50% from a male parent.

B External Validity

B.1 Type of household

The primary goal of our work is to identify gender gaps in households with two parents where one identifies as female and the other as male. About 98% of the US identify as male either male or female, with the remaining 2% identifying in a number of different ways.[1] The plurality of households with children under the age 18, 84%, live in a home with two parents (with 99% of these being opposite gender couples), however 16% have a single caregiver present.[2]

We believe the direction of our effects of our high/low availability messages would be the same for a variety of genders (e.g. two non-binary parents, same sex couples), however we would expect baseline inequality to be closer to zero in these cases. And indeed nationally representative data indicates that same-sex households do not report wishing they were contacted more or less than they actually are by their child's school.[3]

B.2 School Setting

Our experiment takes place in a K-12 school setting. A large portion of the general population, about 40% of households in the US, have school-aged children (NCES, 2021). Schools are an ideal setting to explore external demands on parents' time because 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, volunteering for the book fair or joining the Parent Teacher Association for

[1] `https://www.census.gov/library/stories/2021/11/census-bureau-survey-explores-sexual-orientation-`
`html`

[2] `https://www.census.gov/library/stories/2022/07/most-kids-with-parent-in-same-sex-relationship-li`
`html`

[3] See `https://csed.byu.edu/american-family-survey` for evidence from 205 respondents who are nationally representative. The limited survey evidence we have on non-binary parents from this survey does indicate that the three non-binary respondents report being contacted 75% but wishing to be contacted only 67% of the time.

several reasons. First, educators in our survey report that they would favor contacting the mother first in many of these scenarios (we discuss the survey in Section F).

Second, the gender distribution of these tasks is significantly skewed with mothers comprising almost 90% of Parent Teacher Association members and only 13% of fathers reporting high levels of involvement in their child's school activities, compared to 53% of mothers.[4]

Furthermore, although the gender inequality that we document is in the school setting, this is only one of many settings where mothers spend significantly more time on children than fathers. Prior studies have documented substantial gender differences in time devoted to caring for sick children, taking children to the doctor, and coordinating a wide range of household and child-related tasks, known as cognitive labor.[5] If these other inequalities are also partially driven by external demands, our findings likely represent a lower bound for the overall gender gaps in external demand for parental involvement.

[4]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

[5]Wikle and Cullen (2021); Bianchi et al. (2006); Boye (2015); Daly and Groes (2017); Daminger (2019); Bertrand et al. (2015b); Charmes (2019) https://www.bls.gov/spotlight/2017/differences-in-parents-time-use-between-the-summer-and-the-school-year/home.htm https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0009/3963249/HILDA-Statistical-Report-2021.pdf

C Ethics

A common critique of audit studies which perform outreach from fictitious persons to a third party (often a business that is hiring) is that the person who receives the message wastes time and effort on evaluating the message. We estimate that the median time spent leaving our parents a message was 50 seconds with the 99th percentile being a message of less than two-minutes. As such, each principal in our data set is not spending a large amount of time being in our study. Furthermore, unlike a resume audit study, the principals in our study do not need to evaluate a lengthy resume being considered for a position; rather they only need to read our brief email message and return our call (only 20% of principals call us, and only 17% leave a voicemail further reducing the likelihood of significant harm to our subjects).

We considered writing positive reviews for schools as a form of compensation for their time but after consultation with our IRB were told this would likely be a violation of the terms of service of the review websites, and as such could not gain IRB approval for this.

Also, our subjects are school officials who as part of their position aim to provide increased school quality. Our research in part informs ways to increase school quality through better serving parents, as such participation in our study is partially part of our subjects regular job duties.

A second concern is that the decision-makers in an audit study may harm other non-fictitious persons because of their involvement in the audit study. For example, if a firm decides to callback a fictitious applicant from an audit study this may crowd out a call to real applicant. We believe our study does not pose this same potential harm. The act of calling one family likely does not crowd out further actions. An additional possible harm in a labor market audit study if fictitious applicants with foreign sounding names always decline an interview once it has been extended 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 our households are two-parent households with racially neutral names, as such it is difficult to think about which sub-group a school principal would negatively update about in response to our study.

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, Samek, and van de Ven, 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. The study of gender discrimination is a case where it is often difficult to think of a way to obtain informed consent or a way to use non-fictitious persons without possibly biasing the results. Our study was approved by the relevant Institutional Review Boards at our home institutions, and as such the harms and benefits have been evaluated and approved by a third party.

D Balance Tables

See Tables J5 J6 and J7 for balance in the other Variations of our experiment. See Table K13 for balance on observables when we do not re-weight to account for imbalance in emails sent from male versus female parents.

Table D3: **Balance on Observable Attributes of Schools/Decision Makers By Treatment In Main Variation**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
Elementary	0.48	0.49	0.51	0.50	0.50
Middle	0.14	0.14	0.14	0.15	0.15
High	0.19	0.20	0.20	0.19	0.20
Decison-Maker Female	0.57	0.58	0.59	0.59	0.58
PublicCharter	0.06	0.05	0.06	0.06	0.06
PublicNOTCharter	0.76	0.79	0.81	0.79	0.80
Private	0.18	0.16	0.13	0.15	0.14
FreeLunch	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 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.

E Data Collection & Matching

E.1 Emails and Phone Numbers

To record phone meta-data and voicemails we used a service called Callfire to set up a series of different phone numbers for our male and female parents. We set up the phone numbers with a generic voicemail box and an auto-reply if someone sent a text saying that number did not receive text messages. We set up our email addresses with an auto-reply stating to please call instead of emailing. The exact email addresses that we sent our messages from were "erica@miller-family.net" and "roy@miller-family.net" for part of our data collection. For emails sent during the bulk of data collection, the exact emails were "audrey@the-johnsonfamily.net" and "curtis@the-johnsonfamily.net." Due to some constraints on email send limits, the followup emails sent about two weeks 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." We discuss the choice of exact names in detail below and in Section E.4.

Email is a common way for parents to contact schools. Our own survey of educators found that three-fourths of them report being contacted by parents via email at least once a month, Section F). These educators also reported that when being emailed by both parents, a single parent emailing and CCing 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 E.4), and as such decided to not use a joint family email. Furthermore, we were concerned that a joint family email address might signal a more egalitarian family, and as such might bias our results towards finding more equal calls to mothers and fathers. However, this is one place our analysis is different from our original preregistration where we did not list a plan to analyse results by whether the primary sender of the email was the mother or the father (as we had originally thought we would send emails from a joint family email account).[6]

[6] `https://www.socialscienceregistry.org/trials/7610`

E.2 Names

We chose the names from the top 200 names listed by the social security administration in 1980.[7] We chose 1980 because we primarily contact schools which enroll children between the ages of 5 to 18, so the average age is 11.5 years old. A child who is 11.5 years old now was born in 2009 (2021-11.5=2009.5). The average age of a first time parent in 2009 was 29.4 years old.[8] Our parents on average should have been born in 1980 (because 2009-29.4=1979.6). Within these names we chose names that did not have a strong indication of a specific race or ethnicity (Tzioumis, 2018). We chose our last names (Johnson and Miller) from the list of the most common last names in the US over many decades.[9]. 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 we should be concerned about.[10] In addition we did a google image search for these names to ensure a balance of race and ethnicities was associated with these names.

E.3 Messages

We pre-tested our messages using a survey run on Mechanical Turk to test which messages gave the widest variation in self-reported likelihood of getting a call back. We also pre-tested our messages on a set of educators (see Section F) to ensure they seemed natural to this audience.

Furthermore, for the two message variations we sent the most of (Main and Equal Decision), we tested different versions of the message. The messages we sent were brief by design to use up less of the decision-maker's time and make our treatments that vary information about which parent is more or less available more salient. We did test longer versions of our two most emailed messages as detailed in Table E4. We found that the callback rate was not statistically significantly different between the longer and shorter messages, nor was the proportion of calls to mothers versus fathers.

[7] `https://www.ssa.gov/OACT/babynames/decades/names1980s.html`

[8] CBS News "Average age of first-time mothers up to 29.9 years," November 5, 2019 `https://www.cbs.nl/en-gb/news/2019/19/average-age-of-first-time-mothers-up-to-29-9-years`

[9] `https://namecensus.com/last-names/`

[10] For example, if you google "John List" versus "John List Economist," you will find that there is an infamous American murderer named John List who has way fewer citations than the John List most economists know.

Table E4: Longer Versions of Messages

Variation & Treatment	Body Text
Main Baseline (Used In Study)	We are searching for schools for our child. Can you call one of us to discuss?
Main Baseline (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 (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 (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,

E.4 Pilot Studies

In May of 2021 we sent 767 emails, in June 2021 we sent out 1250 emails, and in Nov 2021 we sent out 1250 emails. The primary purpose of this early data collection was to refine the process by which we send emails, learn about response rates and our ability to match phone calls to emails sent. As such we concentrated on a subset of our treatments: baseline, father high availability, father low availability in the May and June 2021 waves, and expanded to five treatments in the November 2021 wave.

Within our pilots we tested a number of procedural items. In our May pilot we chose the the names Jennifer and Michael because they signal gender well. However, Jennifer and Michael are predominantly white names, so we wanted to test the use of a more race neutral set of names. We tested if the exact first names used had significant differences on callbacks (we tested Jennifer/Michael vs. Erica/Roy) and found that using the more race neutral names decreased callbacks by 8.8 percentage points. We felt 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, our survey of educators found that the use of a joint family email address was less common than individual family email addresses and CCing the other parent (Section F). When we tested these two types of email addresses we found that using a family email address versus individual emails addresses (with one parent CCing the other parent) decreased our callback rates by 9.2 percentage points ($p = 0.032$). With the evidence from the pilot and the survey we decided to move away from a joint family email address in our full data collection efforts.

E.5 Phone Call Data

E.5.1 May 2022 Phone Calls

In May of 2022 we sent out 7,935 emails to a school that had a unique email and unique phone number. We initially sent out more than 7,935 emails in May 2022 for two reasons. First, we did not realize that some schools share a single email or a single phone number (e.g. a network of charter schools or a school district may share a single email or central phone number). Second, an error in our code meant we mistakenly sent some email addresses more than one email. We have removed all these from our data set.

In the weeks following we received 2990 calls our May 2022 emails. In May, some of these calls will be in response to emails we dropped from our data set for the reasons outlined above. Furthermore, these calls include spam calls made to the numbers, but those are most likely randomly distributed across our phone numbers. More of an issue is that these calls include calls made from the same school principal from multiple different phone numbers. They also include calls from the same principal to the same household multiple times in a row, to both the mother and the father, or some combination of this. Our outcome variable of interest is the first parent to be contacted, rather than the total number of calls made by a principal (although this could be of interest also). Furthermore to be able to perform analysis about school or principle specific demographics we need to link each phone call back to a specific email sent. This matching is a multi-step process.

E.5.2 July 2022 Phone Calls

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

E.5.3 Matching Phone Calls To Emails

First, we create a data set with a single line for each unique phone number. We also include all the phone calls from "Restricted" phone numbers as it is impossible to

tell if those are unique. In May 2022 the one-call data set had 1,684 lines, and in June/July 2022 the one-call data set had 17,139 lines. We then match these Callfire 10 digit phone numbers to the 10 digit phone numbers associated with our schools. A little over 60% of calls are matched on 10-digit matches.

Second, we take the remaining Callfire phone calls and perform a "fuzzy" match on the first 6 digits of the phone number. For example, all calls originating from Tufts University start with these same 6 digits 617-627, while 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.

Third, we took any remaining Callfire phone calls and 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 sub-set of these matches to be audited by a different research assistant to check for the quality of our matching.

F Survey of Educators

In April 2022 we ran a survey of educators using the firm Prolific (IRB number STUDY00002608). 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. We had 238 respondents.

One goal of our 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 CCing 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 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 main baseline only 8% of the time, this is very different than the rate we observe in our natural field experiment which is well above 70% not calling back either parent. This could be because some of 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.

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 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.

F.1 Household Survey

Within this survey of educators we also identified which respondents were parents from a household with one male and one female parent. These 91 respondents answered a series of questions about households and schools for us. When asked "What proportion of the time does your child's school contact you versus your partner?" female parents report being contacted 79% of the time while male parents are contacted 41% of the time (note this sums to more than 100%, so each group may incorrectly perceive the reality of who the school contacts more). Interestingly when asked how often they wish they were contacted by the child's school female parents report wanting to be contacted less at 70%, while male parents want to be contacted more at 45%.

Our small sample which is not representative of the US as a whole has fairly similar results to a nationally representative survey which finds that in two-parent heterosexual households with school age children that mothers report being contacted 71% of the time, while males are only contacted 48% of the time.[11] Mothers wish they were contacted less (65% ideally) while fathers in the national sample report wanting to be contacted almost exactly as much as they are contated (at 47% of the time).

We also asked "When the school contacts your family, what proportion of the time do you respond first?" and found female parents report contacting the school first 82% of time while male parents report making first contact only 49% of time. When asked what they wished were the levels of first contact made by each parent male parents said they wished to make first contact 51% of time (similar to the level they report making contact), while female parents wanted to lower their rate of first contact to 76%.

[11] `https://csed.byu.edu/american-family-survey`

G Theory Appendix

G.1 Notation

We provide a summary of our notation as a reference.

Sub- and superscripts

- $i \in I$: decision-maker subscript
- $j \in \{n, f, m\}$: subscript for which parent to call first
- $t \in \{\boldsymbol{b}aseline, \boldsymbol{h}igh\boldsymbol{F}emale, \boldsymbol{l}ow\boldsymbol{F}emale, \boldsymbol{h}igh\boldsymbol{M}ale, \boldsymbol{l}ow\boldsymbol{M}ale\}$: 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 $\boldsymbol{F}$
- $g \in \{\phi, \mu\}$: second superscript for principal gender

Objects of interest

1. Structural parameters: $\delta, r, \omega^2, \lambda$
    - e.g. δ_m^μ for distaste of male principal for calling male parent
2. Reduced form parameters: α, η, γ
    - e.g. $\gamma_m^{\boldsymbol{F},\,\mu}$ for impact of signal about female parent on probability that male principal will call male parents
3. Reduced-form regressors: w, x do not vary with principal gender, so we have w_i^t, x_i^t
4. Reduced-form principal-gender dummy variables: ϕ_i for female principal dummy, $(1 - \phi_i)$ for male principal dummy
5. Proportions of decision-makers: $p_m^{\boldsymbol{F},\,\mu}$
6. Coefficients in treatment effects regression: $\beta^{l\boldsymbol{M}}$, $\beta^{h\boldsymbol{M}}$, $\beta^{l\boldsymbol{F}}$, $\beta^{h\boldsymbol{F}}$
    - e.g. $\beta_m^{l\boldsymbol{F},\,\mu}$ for impact of low signal about female parent on the probability that the male principal will call the male parent

G.2 Base Theoretical Framework

G.2.1 Proof of Result 1 (Identification of Reduced Form Parameters)

In Sections 2.1–2.3, we assume the following:

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, $r_{ij} \sim \mathcal{N}(\bar{r}_j, \omega_j^2)$.
3. $r_{i,f}$ and $r_{i,m}$ are independent.
4. Decision makers are risk neutral.
5. 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.
6. Each decision maker has a distaste for calling that varies by alternative. Decision maker i has a taste for calling j when $\delta_{ij} < 0$.
7. Each decision maker i knows c_i and δ_{ij}.
8. Expected utility for decision maker i is $\mathbb{E}(U_{ij}) = \mathbb{E}(r_{ij}) - (\delta_{ij} + c_i)$ for $j \in \{n, f, m\}$ with $\mathbb{E}(U_{ij}) = 0$.
9. The experimenters choose signal values x_{ij} at random to show each decision maker and send a signal $x_{ij} \in \{-1, 1\}$ about at most one alternative to each decision maker. The decision makers believe that $x_{ij} \sim \mathcal{N}(r_j, \sigma^2),\ j \in \{f, m\}$, where r_j is the true responsiveness of j.
10. A signal x_{ij} can shift the belief $\tilde{r}_{ij}$ but does not affect c_i or δ_{ij}.
11. ε_{ij} are each distributed according to the standard Gumbel distribution.

Given the above assumptions and the experimental data, we can use the observable proportions of decision makers in each signal-outcome pair to identify the reduced-form parameters.

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 and γ_j are zero for all decision makers, so

we have $U_{ij} = \alpha_j \ \forall j$. Because $U_{i,n} = \alpha_n = 0$ by assumption, the probabilities from the logit model are

$$p_n^b \equiv \frac{1}{1 + e^{\alpha_f} + e^{\alpha_m}} \quad p_f^b \equiv \frac{e^{\alpha_f}}{1 + e^{\alpha_f} + e^{\alpha_m}} \quad p_m^b \equiv \frac{e^{\alpha_m}}{1 + e^{\alpha_f} + e^{\alpha_m}}$$

where subscripts denote which alternative is chosen and the superscript "0" denotes that no signal is sent about either alternative.

Sending a signal ($w_{i,f} = 1$) with value $x_{i,f} = 1$ about alternative f and no signal about alternative m makes the deterministic part of utility for alternative f (i.e. Equation 4 without the error) $\alpha_f + \eta_f + \gamma_f$. We therefore have the following probabilities:

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

where the superscript "$h\boldsymbol{F}$" denotes that we send only a high signal (i.e. value of 1) about alternative f.

Similarly, when we send a signal with value $x_{i,f} = -1$ about alternative f and no signal about alternative m makes the deterministic part of utility for alternative f (i.e. Equation 4 without the error) $\alpha_f + \eta_f - \gamma_f$. We therefore have the following probabilities:

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

where 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 ($w_{i,m} = 1$) with value $x_{i,m} = 1$ about alternative m and no signal about alternative f leads to the following probabilities:

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

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

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

Next, we manipulate the logit probabilities to identify reduced-form parameters $\alpha_j, \eta_j, \gamma_j$. In order to identify α_j, we take ratios of the probabilities for when no signal is sent.

$$\frac{p_j^b}{p_n^b} = e^{\alpha_j} \Leftrightarrow \boxed{\alpha_j = \ln p_j^b - \ln p_n^b \text{ for } j \in \{f, m\}} \tag{11}$$

In order to identify η_j and γ_j, we must combine equations. Start with

$$\frac{p_j^{hJ}}{p_n^{hJ}} = e^{\alpha_j + \eta_j + \gamma_j} \Leftrightarrow \alpha_j + \eta_j + \gamma_j = \ln p_j^{hJ} - \ln p_n^{hJ} \tag{12}$$

and

$$\frac{p_j^{lJ}}{p_n^{lJ}} = e^{\alpha_j + \eta_j - \gamma_j} \Leftrightarrow \alpha_j + \eta_j - \gamma_j = \ln p_j^{lJ} - \ln p_n^{lJ} \tag{13}$$

where $J \in \{F, M\}$ denotes about which parent the signal is sent.

Subtracting Equation (13) from Equation (12), we have

$$\alpha_j + \eta_j + \gamma_j - \alpha_j - \eta_j + \gamma_j = \ln p_j^{hJ} - \ln p_n^{hJ} - \ln p_j^{lJ} + \ln p_n^{lJ}$$

Simplifying, we have

$$2\gamma_j = \ln p_j^{hJ} - \ln p_n^{hJ} - \ln p_j^{lJ} + \ln p_n^{lJ} \Leftrightarrow \boxed{\gamma_j = \tfrac{1}{2}\left[\ln p_j^{hJ} - \ln p_n^{hJ} - \ln p_j^{lJ} + \ln p_n^{lJ}\right] \text{ for } j \in \{f, m\}} \tag{14}$$

Combining Equations (11), (12) and (13), we have

$$\ln p_j^b - \ln p_n^b + \eta_j + \frac{1}{2}\left[\ln p_j^{hJ} - \ln p_n^{hJ} - \ln p_j^{lJ} + \ln p_n^{lJ}\right] = \ln p_j^{hJ} - \ln p_n^{hJ}$$

Simplifying

$$\eta_j = -\ln p_j^b + \ln p_n^b + \ln p_j^{hJ} - \ln p_n^{hJ} - \frac{1}{2}\left[\ln p_j^{hJ} - \ln p_n^{hJ} - \ln p_j^{lJ} + \ln p_n^{lJ}\right]$$

$$\boxed{\eta_j = -\ln p_j^b + \ln p_n^b + \tfrac{1}{2}\left[\ln p_j^{hJ} - \ln p_n^{hJ} + \ln p_j^{lJ} - \ln p_n^{lJ}\right] \text{ for } j \in \{f, m\}} \tag{15}$$

■

It is worth noting that η_j and γ_j cannot vary independently: every term in Expression (14) is also present in Expression (15).

G.2.2 Proof of Result 2 (Identification of Structural Parameters)

We use the six identified reduced-form parameters from Result 1 to identify the deep parameters $\lambda_f, \lambda_m, \bar{r}_f, \bar{r}_m$ and $\bar{\delta}_f - \bar{\delta}_m$. Recall Equations (5)-(7) that relate the reduced-form parameters to the deep parameters:

$$\alpha_j = \bar{r}_j - \bar{\delta}_j - c$$
$$\eta_j = -(1 - \lambda_j)\bar{r}_j$$
$$\gamma_j = 1 - \lambda_j$$

We can directly identify $\lambda_j = 1 - \gamma_j$ using the third of these equations. Recall that λ_j is composed of σ^2 and ω_j^2, but these can't be separately identified since we do not vary ω_j^2.

We can then combine the equation for λ_j with the second equation to find $\bar{r}_j = -\frac{\eta_j}{\gamma_j}$. Plugging this into the first equation produces $\bar{\delta}_j + c = -\frac{\eta_j}{\gamma_j} - \alpha_j$.

We cannot separately identify $\bar{\delta}_f$ and $\bar{\delta}_m$, but we can combine $\bar{\delta}_f + c = -\frac{\eta_f}{\gamma_f} - \alpha_f$ and $\bar{\delta}_m + c = -\frac{\eta_m}{\gamma_m} - \alpha_m$ from the previous step to get $\bar{\delta}_m - \bar{\delta}_f = \frac{\eta_f}{\gamma_f} - \frac{\eta_m}{\gamma_m} + \alpha_f - \alpha_m$. ■

G.2.3 Additional Testable Hypotheses

We add the following, more detailed testable hypotheses to those in Section 2.4.

Hypothesis 4. *When decision-makers receive a signal that the male parent has a lot of availability (treatment h**M**), the proportion who call the female parent first will be smaller than the proportion of decision makers who receive no signal and call the female parent first, i.e.* $p_f^{h\boldsymbol{M}} < p_f^b$. *Likewise, when decision makers receive a signal that the female parent has a lot of availability (treatment h**F**), the proportion who call the female parent first will be larger than the proportion who receive no signal and call the female parent first, i.e.* $p_f^b < p_f^{h\boldsymbol{F}}$.

Hypothesis 5. *The male high availability signal will have a larger impact than the female high availability signal on the proportion of decision makers who call the female parent first, i.e.* $p_f^{h\boldsymbol{F}} - p_f^b < p_f^b - p_f^{h\boldsymbol{M}}$.

We interpret support for this hypothesis to indicate that messages that contradict the decision-makers' priors (here, that female parents are more available) have a larger impact than messages that confirm their priors. We can write three more versions of the inequality in Hypothesis 5: one for the impact of the positive signals on calling the male parent, with the remaining two comparing the impacts of the negative signals.

G.3 Model with Decision-Maker Characteristics

We let g index the discrete categories that make up the decision-maker characteristic. Each type g of the decision-maker makes their decision as in Section 2.3. This model extension applies to any observable characteristic of decision makers that is discrete in nature. Here, we focus on the gender of the decision maker so that $G = \{\phi, \mu\}$, where female decision makers are denoted by ϕ and male decision makers denoted by μ.

With decision-maker characteristics, Equation 1 becomes

$$\mathbb{E}(U_{ij}^g) = \mathbb{E}(r_{ij}^g) - \delta_{ij}^g - c_i$$

Each type g of the decision-maker makes their decision as in Section 2.3. 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 2.3 so that beliefs are not only independent across alternatives but also across types of decision maker, i.e. that all $r_{ij}^g \sim \mathcal{N}(\bar{r}_j^g, \omega_j^2)$ are mutually independent.

We now have that decision maker i of gender g has a posterior mean $\widetilde{r}_{ij}^g$ for the responsiveness of j, assuming the prior variance is *common* to all i, is

$$\widetilde{r}_{ij}^g = \lambda_j^g \bar{r}_{ij}^g + (1 - \lambda_j^g) x_{ij}, \quad \lambda_j^g = \frac{1/(\omega_j^g)^2}{1/(\omega_j^g)^2 + 1/\sigma^2}.$$

We continue to assume that decision maker beliefs are not heterogeneous within type, so that $\bar{r}_{ij}^g = \bar{r}_j^g \ \forall i$. Since signals are not differentiated by decision-maker type, Equation 3 becomes

$$\mathbb{E}(U_{ij}^g) = \bar{r}_j^g - (1 - \lambda_j^g)\bar{r}_j^g w_{ij} + (1 - \lambda_j^g) w_{ij} x_{ij} - \delta_{ij}^g - c_i$$

for all j and g.

Equations (4)-(8) become

$$U_{ij}^g = \alpha_j^g + \eta_j^g w_{ij} + \gamma_j^g w_{ij} x_{ij} + \varepsilon_{ij}^g \tag{16}$$

$$\alpha_j^g = \bar{r}_j^g - \bar{\delta}_j^g - c \tag{17}$$

$$\eta_j^g = -(1 - \lambda_j^g)\bar{r}_j^g \tag{18}$$

$$\gamma_j^g = 1 - \lambda_j^g \tag{19}$$

$$\varepsilon_{ij}^g = (c - c_i) + (\bar{\delta}_j^g - \delta_{ij}^g) \tag{20}$$

where $\bar{\delta}_j^g$ denotes the average value of δ_{ij}^g.

We then have the following identification result:

Result 4. *Given the assumptions of Sections 2.1–2.4 and this section, the reduced-form parameters α_j^g, γ_j^g, η_j^g and the structural parameters $\lambda_f^g, \lambda_m^g, \bar{r}_f^g, \bar{r}_m^g$, $\bar{\delta}_f^g - \bar{\delta}_m^g$ are identified for $j \in \{f, m\}$ and $g \in G$, G discrete.*

Proof: Repeatedly apply the proofs for Results 1 and 2 for each $g \in G$. ■

G.3.1 Testable Hypotheses

Result 4 allows us to identify beliefs and preferences for each gender of decision maker. All the testable hypotheses from Sections 2.4 and G.2.3 can again be tested here, with one version for each characteristic of the decision makers. The following hypotheses are of particular interest as regards the theory of homophily.

Hypothesis 6. *Decision-makers display homophily, meaning they are more likely to call a parent of their same gender. That means that, when no signal is sent about availability, the proportion of calls to male parents will be higher when the decision maker is male, and that the proportion of calls to female parents will be higher when the decision maker is female. We find support for this hypothesis if $p_m^{b,\mu} > p_f^{b,\mu}$ and $p_f^{b,\phi} > p_m^{b,\phi}$ where decision maker's gender is represented by superscripts ϕ and μ.*

This is equivalent to $\alpha_m^\mu > \alpha_f^\mu$ and $\alpha_f^\phi > \alpha_m^\phi$ in terms of the reduced-form parameters. In terms of the structural parameters, it is $\bar{r}_m^\mu - \bar{\delta}_m^\mu > \bar{r}_f^\mu - \bar{\delta}_f^\mu$ and $\bar{r}_f^\phi - \bar{\delta}_f^\phi > \bar{r}_m^\phi - \bar{\delta}_m^\phi$.

Hypothesis 7. *The relative distaste for calling male parents versus female parents is larger for female decision makers than it is for male decision makers. We find support for this hypothesis if* $\bar{\delta}^{\phi}_{m} - \bar{\delta}^{\phi}_{f} > \bar{\delta}^{\mu}_{m} - \bar{\delta}^{\mu}_{f}$.

This hypothesis derives directly from the definition of homophily, that is that male decision makers have less distaste for female parents than they do for male parents ($\bar{\delta}^{\mu}_{m} < \bar{\delta}^{\mu}_{f}$) and female decision makers have less distaste for male parents than they do for female parents ($\bar{\delta}^{\phi}_{f} < \bar{\delta}^{\phi}_{m}$). Combining these two inequalities provides the condition in terms of identifiable parameters, that is, the differences between distaste parameters for each type of decision maker.

G.4 Relaxing the independence assumption

We now assume (1) the distributions of the signals about the two parents can not only have different means but also different variances and (2) the impact on decision-maker updating can be summarized by a correlation parameter ρ_j, which captures the impact on the belief about parent j from a signal about the other parent.

In addition to the new structural parameters ρ_j, this version of the model also has update parameters λ^t_j that are differentiated not only by the parent about whom the update is being made (j), but now also by the parent about whom the message is sent (t). The reduced form parameters η and γ are also now differentiated by the parent about whom the signal is sent.

After relaxing the assumption that a signal about one parent only affects the belief about that parent, the updating process becomes more complex. Note that, in order to keep notation simple, we focus without loss of generality on how the belief about the female parent is updated. Equation (2) becomes

$$\widetilde{r}^F_{if} = \lambda^F_f \bar{r}_f + (1 - \lambda^F_f) x_{if}, \quad \lambda^F_f = \frac{1/\omega^2_f}{1/\omega^2_f + 1/\sigma^2_F} \tag{21}$$

$$\widetilde{r}^M_{if} = \lambda^M_f \bar{r}_f + (1 - \lambda^M_f) \rho_f x_{im}, \quad \lambda^M_f = \frac{1/\omega^2_f}{1/\omega^2_f + 1/\sigma^2_M} \tag{22}$$

with

$$\bar{r}_f \sim \mathcal{N}(r_f, \omega^2_f), \quad x_{if} \sim \mathcal{N}(r_f, \sigma^2_F), \quad x_{im} \sim \mathcal{N}(r_m, \sigma^2_M).$$

We now have two ways that decision maker i's belief about the female parent can be

updated: via a signal directly about the female parent ($w_{i,f} = 1$) and $w_{i,m} = 0$), or via a signal about the male parent ($w_{i,f} = 0$) and $w_{i,m} = 1$).[12]

Under this more general formulation, Equation (3) becomes

$$\mathbb{E}(U_{i,f}) = (1 - w_{i,f} - w_{i,m})\bar{r}_f + w_{i,f}\widetilde{r}^F_{i,f}(x_{i,f}, x_{i,m}) + w_{i,m}\widetilde{r}^M_{i,f}(x_{i,f}, x_{i,m}) - (\delta_{ij} + c_i) \tag{23}$$

$$= (1 - w_{i,f} - w_{i,m})\bar{r}_f + w_{i,f}\left[\lambda^F_f \bar{r}_f + (1 - \lambda^F_f)x_{i,f}\right] + \tag{24}$$

$$w_{i,m}\left[\lambda^M_f \bar{r}_f + (1 - \lambda^M_f)(\rho_f)x_{i,m}\right] - (\delta_{ij} + c_i) \tag{25}$$

$$= \bar{r}_j - (1 - \lambda^F_f)\bar{r}_f w_{i,f} - (1 - \lambda^M_f)\bar{r}_f w_{i,m} + \tag{26}$$

$$(1 - \lambda^F_f)w_{i,f}x_{i,f} + (1 - \lambda^M_f)\rho_f w_{i,m}x_{i,m} - (\delta_{ij} + c_i) \tag{27}$$

$$= \alpha_f + \eta^F_f w_{i,f} + \eta^M_f w_{i,m} + \gamma^F_f w_{i,f}x_{i,f} + \gamma^M_f w_{i,m}x_{i,m} + \varepsilon_{i,f}. \tag{28}$$

where the last equation follows from the mapping below:

$$\begin{aligned}
\alpha_f &= \bar{r}_f - \bar{\delta}_f - c \\
\eta^F_f &= -(1 - \lambda^F_f)\bar{r}_f \\
\eta^M_f &= -(1 - \lambda^M_f)\bar{r}_f \\
\gamma^F_f &= 1 - \lambda^F_f \\
\gamma^M_f &= (1 - \lambda^M_f)\rho_f \\
\varepsilon_{i,f} &= (c - c_i) + (\bar{\delta}_f - \delta_{i,f}).
\end{aligned}$$

Result 5. *Given the assumptions of Sections 2.1–2.5.3, the reduced-form parameters α_j, γ^t_j, η^t_j and the structural parameters $\lambda^t_j, \rho_j, \bar{r}_j$ and $\bar{\delta}_f - \bar{\delta}_m$ are identified for $j \in \{f, m\}$ and $t \in \{F, M\}$.*

Proof: α_f is once again identified by Equation (5). Equations (6) and (7) identify η^F_f and γ^F_f (with only a notational change from η_f and γ_f to η^F_f and γ^F_f). The following equations identify η^M_f and γ^M_f.

[12] This formulation is relatively simple because we only send signals about one parent to any given decision maker. It can be generalized for the case where one sends signals about both parents to the same decision maker.

$$\frac{p_f^{hM}}{p_n^{hM}} = e^{\alpha_f+\eta_f^M+\gamma_f^M\rho_f} \Leftrightarrow \alpha_f+\eta_f^M+\gamma_f^M\rho_f = \ln\left(\frac{p_f^{hM}}{p_n^{hM}}\right) \Leftrightarrow \alpha_f+\eta_f^M+\gamma_f^M\rho_f = \ln p_f^{hM}-\ln p_n^{hM} \tag{29}$$

$$\frac{p_f^{lM}}{p_n^{lM}} = e^{\alpha_f+\eta_f^M-\gamma_f^M\rho_f} \Leftrightarrow \alpha_f+\eta_f^M-\gamma_f^M\rho_f = \ln\left(\frac{p_f^{lM}}{p_n^{lM}}\right) \Leftrightarrow \alpha_f+\eta_f^M-\gamma_f^M\rho_f = \ln p_f^{lM}-\ln p_n^{lM} \tag{30}$$

Subtracting Equation (30) from Equation (29), we have

$$\alpha_f + \eta_f^M + \gamma_f^M\rho_f - \alpha_f - \eta_f^M + \gamma_f^M\rho_f = \ln p_f^{hM} - \ln p_n^{hM} - \ln p_f^{lM} + \ln p_n^{lM}$$

Simplifying, we have

$$2\rho_f\gamma_f^M = \ln p_f^{hM}-\ln p_n^{hM}-\ln p_f^{lM}+\ln p_n^{lM} \Leftrightarrow \boxed{\rho_f\gamma_f^M = \tfrac{1}{2}\left[\ln p_f^{hM} - \ln p_n^{hM} - \ln p_f^{lM} + \ln p_n^{lM}\right]} \tag{31}$$

Combining Equations (11), (29) and (30), we have

$$\ln p_f^b - \ln p_n^b + \eta_f^M + \frac{1}{2}\left[\ln p_f^{hM} - \ln p_n^{hM} - \ln p_f^{lM} + \ln p_n^{lM}\right] = \ln p_f^{hm} - \ln p_n^{hM}$$

Simplifying

$$\eta_f^M = -\ln p_f^b + \ln p_n^b + \ln p_f^{hM} - \ln p_n^{hM} - \frac{1}{2}\left[\ln p_f^{hM} - \ln p_n^{hM} - \ln p_f^{lM} + \ln p_n^{lM}\right]$$

$$\boxed{\eta_f^M = -\ln p_f^b + \ln p_n^b + \tfrac{1}{2}\left[\ln p_f^{hM} - \ln p_n^{hM} + \ln p_f^{lM} - \ln p_n^{lM}\right]} \tag{32}$$

Analogous equations similarly identify η_m^F and γ_m^F. Combined with the results for the male parent in the proof of Result 1, all reduced-form parameters in the model generalized for correlations are identified.

It is left to show that the structural parameters are identified. Again, it is without loss of generality to demonstrate identification for the parameters about the female parent; analogous equations for the male parent hold.

As in the proof of Result 2, γ_f^F directly identifies λ_f^F as $\lambda_f^F = 1 - \gamma_f^F$. Once we have λ_f^F, we combine it with the η_f^F equation to get $\bar{r}_f = -\frac{\eta_f^F}{\gamma_f^F}$. Next, we use the η_f^M equation to get $\lambda_f^M = 1 - \frac{\eta_f^M}{\eta_f^F}\gamma_f^F$. We now have everything we need to derive $\rho_f = \frac{\eta_f^F}{\gamma_f^F}$.

$\frac{\gamma_f^M}{\eta_f^M}$ from the γ_f^M equation. Finally, from the α_f equation, we have $\bar{\delta}_f + c = -\frac{\eta_f^F}{\gamma_f^F} - \alpha_f$. Subtracting this equation from $\bar{\delta}_m + c = -\frac{\eta_m^M}{\gamma_m^M} - \alpha_m$, we have $\bar{\delta}_m - \bar{\delta}_f = \frac{\eta_f^F}{\gamma_f^F} - \frac{\eta_m^M}{\gamma_m^M} + \alpha_f - \alpha_m$ as in Result 2. ■

Careful examination of the proof of Result 5 compared to those of Results 1 and 2 will make clear that the identification of the parameters in the base model 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 baseline message. Generalizing the model to allow for this correlation simply lets us test the independence assumption and then to quantify the size of the correlation and any potential differences in the updating processes after signals about male versus female parents.

G.4.1 Testable Hypothesis

Now that we have established that both the reduced-form and structural parameters are well-identified, we put forward an additional testable hypothesis that emerges from the extended model. Note that all the testable hypotheses from Sections 2.4 and G.2.3 remain valid under this model extension.

Hypothesis 8. *Decision makers infer information:*

i) about the male parent after receiving a signal about the female parent; and

ii) about the female parent after receiving a signal about the male parent.

We find support for this hypothesis if (i) $\rho_m \neq 0$ *and (ii)* $\rho_f \neq 0$.

The interpretation of the sign of ρ_j is as follows: If ρ_j is positive, a positive (negative) signal about parent j is taken to also be a positive (negative signal about the other parent. If ρ_j is negative, a positive (negative) signal about parent j is taken to be a *negative (positive)* signal about the other parent.

Note that Hypothesis 4 already establishes that decision makers infer information about the parent on which they receive a signal; Hypothesis 8 says that they also infer information about the other parent.

G.5 Risk Aversion

To develop intuition for the effect of risk aversion, imagine that a decision maker holds the same beliefs and has the same distaste parameter 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.

We can infer the ordering of the prior belief variance by comparing the weights that decision makers place on the prior belief, λ_m and λ_f. Assume without loss of generality that decision makers place greater weight on the prior variance for the female parent, that is $\lambda_f > \lambda_m$. This implies that the prior and posterior variance for the belief about females is lower, i.e. $\omega_f^2 < \omega_m^2$. Intuitively, decision makers place less weight on the prior belief when the prior belief is more uncertain.

H Additional Extensions

H.1 Homophily

When thinking about homophily we are trying to test if the effects are different depending on the pairings of genders between the decision maker and the parent they are contacting. There are four possible pairings (Mm, Ff, Mf, Fm) if assuming all decision-makers and parents are male or female (we know that in real world this is not the case, but this is for simplicity).

Let's consider if we wanted to look for homophily within the baseline alone (that is within decision-makers who were sent the baseline message).

$$Y_{i,j}^* = H_{baseline} X_{ij} + z_i \gamma_j + \alpha_{i,j}$$

$Y_{i,j}$ = j is the choice of called mom, called dad, called no one
We observe:

$$Y_i = argmax(y_{callmom}^*, y_{calldad}^*, y_{nocall}^*)$$

Principal i
X_{ij} are choice / DM specific covariates (e.g. Male DM calls male parent, Female DM calls male parent).

z_i are DM specific covariates (e.g. whether DM has brown eyes)

Model I end up running in Stata is something like

$$Y^*_{i,j} = H_{baseline} X_{ij} + \alpha_{i,j}$$

Where the only co-variate is one that is DM/ choice specific.

Then I can also run a model where I control for the treatment (this assumes the treatment doesn't interact with homophily).

$$Y^*_{i,j} = HX_{ij} + \beta^{l\boldsymbol{M}}(lowMale) + \beta^{h\boldsymbol{M}}(highMale) + \beta^{l\boldsymbol{F}}(lowFemale) + \beta^{h\boldsymbol{F}}(highFemale) + \alpha_{i,j}$$

H.2 Mapping to reduced form parameters for gender-differentiated model

For Pilot 2 without signals about female parents (TO BE UPDATED WHEN DATA IS COMPLETE), we can run the following treatment effects regression:

$$\begin{aligned} y_i = c_f(\phi_i) + c_m(1-\phi_i) + \beta^{l\boldsymbol{M},\mu}(lowMale*(1-\phi_i)) + \beta^{h\boldsymbol{M},\mu}(highMale*(1-\phi_i)) + \\ \beta^{l\boldsymbol{F},\phi}(lowMale*\phi_i) + \beta^{h\boldsymbol{F},\phi}(highMale*\phi_i) + \epsilon_i \end{aligned}$$

Here ϕ_i is a dummy variable that is equal to 1 when i is female. Run an unordered logit where $y_i = 0$ if called neither; $y_i = 1$ if called female; $y_i = 2$ if called male.

In "CallFemale" equation, we get

$$\begin{aligned} c_f^{\mu} &= -0.877 \\ c_f^{\phi} &= -1.076 \\ \beta_f^{l\boldsymbol{M},\mu} &= 0.009 \\ \beta_f^{h\boldsymbol{M},\mu} &= -2.198 \\ \beta_f^{l\boldsymbol{M},\phi} &= 0.253 \\ \beta_f^{h\boldsymbol{M},\phi} &= -0.726 \end{aligned}$$

In "CallMale" equation, we get

$$
\begin{aligned}
c_m^{\mu} &= -1.444 \\
c_f^{\phi} &= -2.194 \\
\beta_m^{l\boldsymbol{M},\mu} &= -0.285 \\
\beta_m^{h\boldsymbol{M},\mu} &= 0.704 \\
\beta_m^{l\boldsymbol{M},\phi} &= -0.088 \\
\beta_m^{h\boldsymbol{M},\phi} &= 1.490
\end{aligned}
$$

The reduced-form regression equation for this model is

$$U_{ij} = \alpha_j^{\phi} * \phi_i + \alpha_j^{\mu} * (1-\phi_i) + \eta_j^{\mu} w_{ij} \cdot (1-\phi_i) + \eta_j^{\phi} w_{ij} \cdot \phi_i + \gamma_j^{\mu} w_{ij} x_{ij} \cdot (1-\phi_i) + \gamma_j^{\phi} w_{ij} x_{ij} \cdot \phi_i + \varepsilon_{ij}.$$

It matters both who you send a message about AND who gets called. We use the same notation as above, where capital superscripts indicate about whom the message is sent, lowercase subscripts indicate who gets called, and principal gender is denoted by a second subscript ϕ or μ. We get

$$
\begin{aligned}
\alpha_f^{\mu} &= -0.877 \\
\alpha_f^{\phi} &= -0.199 \\
\eta_f^{\boldsymbol{M},\mu} &= -1.095 \\
\eta_f^{\boldsymbol{M},\phi} &= -0.237 \\
\gamma_f^{\boldsymbol{M},\mu} &= -1.103 \\
\gamma_f^{\boldsymbol{M},\phi} &= -0.490
\end{aligned}
$$

$$
\begin{aligned}
\alpha_m^{\mu} &= -1.444 \\
\alpha_m^{\phi} &= -0.750 \\
\eta_m^{\boldsymbol{M},\mu} &= 0.210 \\
\eta_m^{\boldsymbol{M},\phi} &= 0.701 \\
\gamma_m^{\boldsymbol{M},\mu} &= 0.494 \\
\gamma_m^{\boldsymbol{M},\phi} &= 0.789
\end{aligned}
$$

Notice that the constants are the same in the treatment effects and reduced-form equations. The relationship between the other sets of parameters is as follows:

- $\eta_m^{\boldsymbol{M},\mu} + \gamma_m^{\boldsymbol{M},\mu} = \beta_m^{h\boldsymbol{M},\mu}$ (signal value of 1 / highMale and male calling Male), i.e. $0.210 + 0.494 = 0.704$
- $\eta_m^{\boldsymbol{M},\mu} - \gamma_m^{\boldsymbol{M},\mu} = \beta_m^{l\boldsymbol{M},\mu}$ (signal value of -1 / lowMale), i.e. $0.210 - 0.494 \approx -0.285$
- $\eta_m^{\boldsymbol{M},\phi} + \gamma_m^{\boldsymbol{M},\phi} = \beta_m^{h\boldsymbol{M},\phi}$ (signal value of 1 / highMale and male calling Male), i.e. $0.701 + 0.789 = 1.490$
- $\eta_m^{\boldsymbol{M},\phi} - \gamma_m^{\boldsymbol{M},\phi} = \beta_m^{l\boldsymbol{M},\phi}$ (signal value of -1 / lowMale), i.e. $0.701 - 0.789 \approx -0.088$
- $\eta_f^{\boldsymbol{M},\mu} + \gamma_f^{\boldsymbol{M},\mu} = \beta_f^{h\boldsymbol{M},\mu}$ (signal value of 1 / highMale & male calling Female) and $\eta_f^{\boldsymbol{M},\mu} - \gamma_f^{\boldsymbol{M},\mu} = \beta_f^{l\boldsymbol{M},\mu}$ (signal value of -1 / lowMale)
- $\eta_f^{\boldsymbol{M},\phi} + \gamma_f^{\boldsymbol{M},\phi} = \beta_f^{h\boldsymbol{M},\phi}$ (signal value of 1 / highMale & female calling Female) and $\eta_f^{\boldsymbol{M},\phi} - \gamma_f^{\boldsymbol{M},\phi} = \beta_f^{l\boldsymbol{M},\phi}$ (signal value of -1 / lowMale)

We can now use Equations 18–20 to derive the a subset of the structural param-

eters from the reduced form parameters, separately for male and female principals:[13]

$$\begin{aligned}
\lambda_m^{\phi} &= 0.211 \\
\lambda_m^{\mu} &= 0.506 \\
\bar{r}_m^{\phi} &= -0.888 \\
\bar{r}_m^{\mu} &= -0.423 \\
\bar{\delta}_m^{\phi} + c &= 1.306 \\
\bar{\delta}_m^{\mu} + c &= 1.021
\end{aligned}$$

We can interpret these results to imply that

1. female principals put less weight on their priors (i.e. they update more strongly toward the signal than male principals do);

2. female principals are even more pessimistic about the availability of male parents than are male principals; and

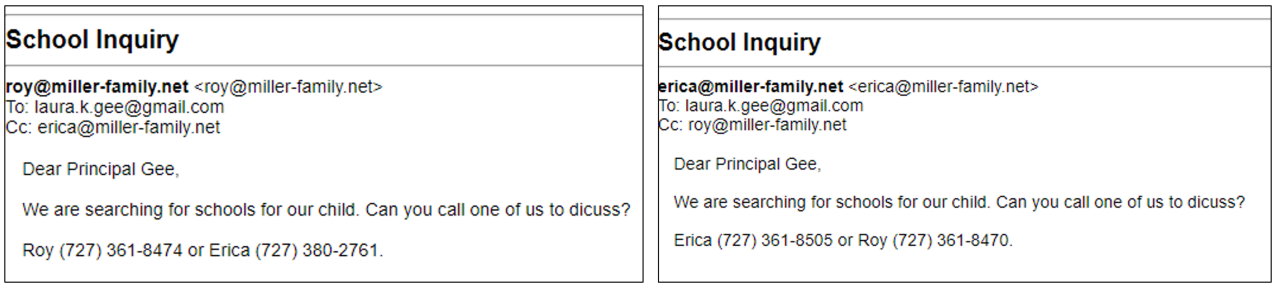
3. Talking about the total cost of calling a male parent (the direct cost of calling combined with the distaste for calling males) is larger for female principals than it is for male principals.

We can combine $\bar{\delta}_m^{\phi}+c$ and $\bar{\delta}_m^{\mu}+c$ to infer the difference between the distaste of female principals and male principals for calling a male parent is $\bar{\delta}_m^{\phi}-\bar{\delta}_m^{\mu} = 0.275$. Because we cannot separately identify the cost of calling, we can only say that female principals have a larger distaste parameter for calling male parents than male principals do. Three things could be going on here; (1) female principals have more distaste for calling male parents; (2) female principals have distaste for calling males while male principals have a taste for calling male parents; or (3) both male and female principals have a taste for calling male parents but that taste is larger for male principals.

I Example Emails Full Text

[13]In order to identify all the parameters of interest, we need to send signals about the female parent as well as the male parent.

Figure I1: Main 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 dicuss?

Roy (727) 361-8474 or Erica (727) 380-2761.

School Inquiry

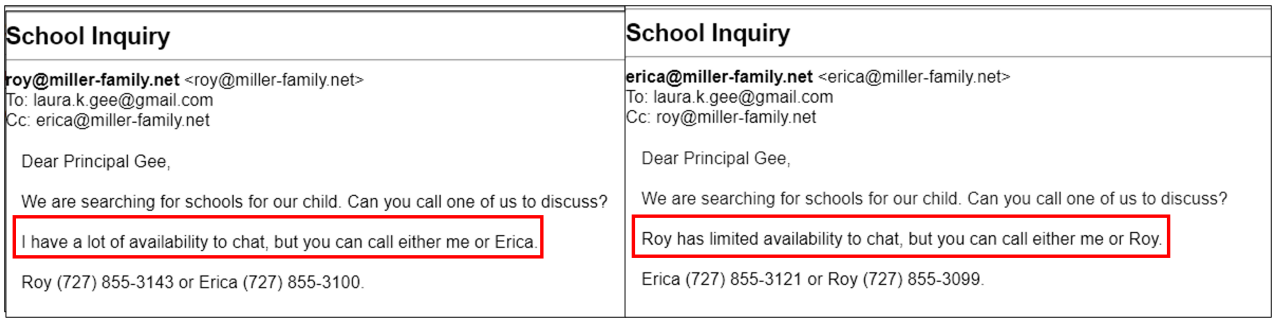
erica@miller-family.net <erica@miller-family.net>
To: laura.k.gee@gmail.com
Cc: roy@miller-family.net

Dear Principal Gee,

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

Erica (727) 361-8505 or Roy (727) 361-8470.

Figure I2: Main: 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

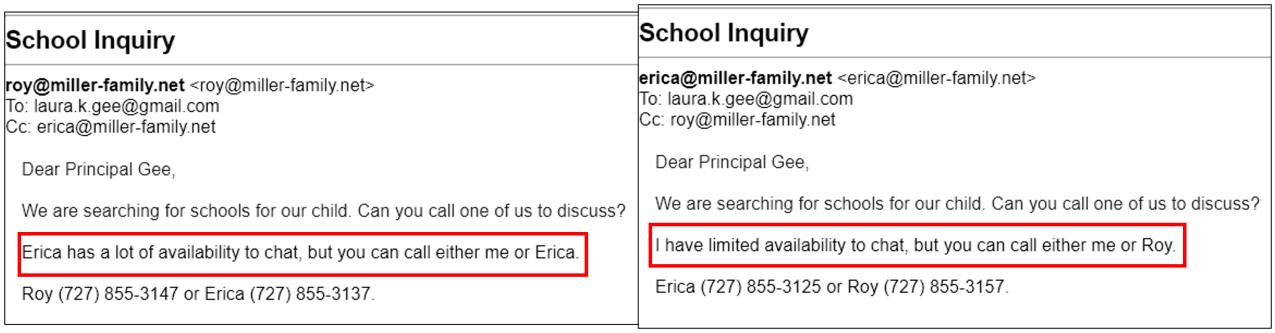
Dear Principal Gee,

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.

Figure I3: Main: 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.

J Variations On Main Messages

J.1 Balance Tables For Variations

Table J5: **Balance on Observable Attributes of Schools/Decision Makers By Treatment In Equal Decision Variation**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
Elementary	0.50	0.50	0.49	0.48	0.48
Middle	0.15	0.15	0.13	0.14	0.14
High	0.20	0.20	0.18	0.19	0.18
Decison-Maker Female	0.58	0.58	0.57	0.58	0.57
PublicCharter	0.06	0.05	0.06	0.06	0.05
PublicNOTCharter	0.80	0.80	0.77	0.76	0.76
Private	0.14	0.14	0.18	0.18	0.18
FreeLunch	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: 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 J6: **Balance on Observable Attributes of Schools/Decision Makers By Treatment In Full Time Variation**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
Elementary	0.49	0.51	0.48	0.52	0.49
Middle	0.17	0.17	0.14	0.16	0.14
High	0.21	0.22	0.18	0.21	0.20
Decison-Maker Female	0.56	0.59	0.57	0.60	0.59
PublicCharter	0.06	0.06	0.05	0.06	0.05
PublicNOTCharter	0.80	0.82	0.73	0.81	0.77
Private	0.14	0.12	0.22	0.13	0.18
FreeLunch	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

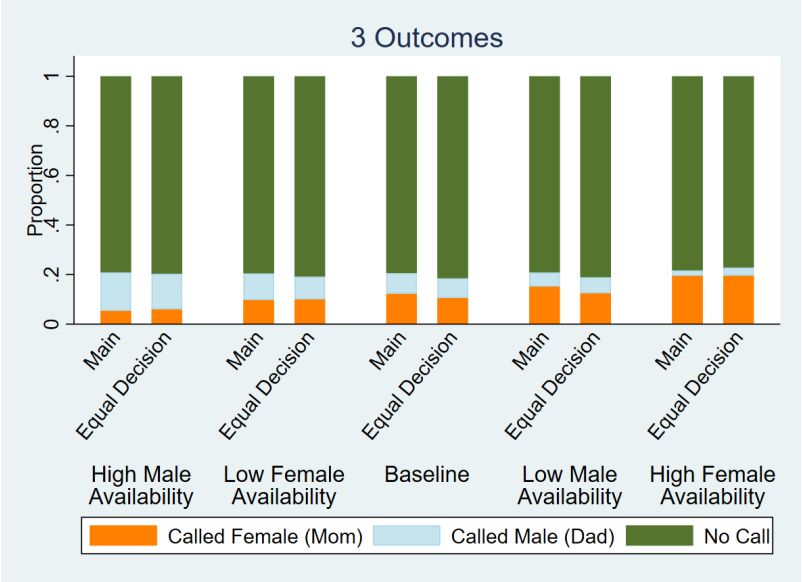
Notes: There is a small proportion of schools which are not Elementary, Middle or High Schools (e.g. K-12 or pre-schools). The following variables are only known for non-private schools: FreeLunch, White, Black, Hispanic. DMFemale is whether the decision-maker (the principal) has a first name that is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

Table J7: **Balance on Observable Attributes of Schools/Decision Makers By Treatment In Payments Variation**

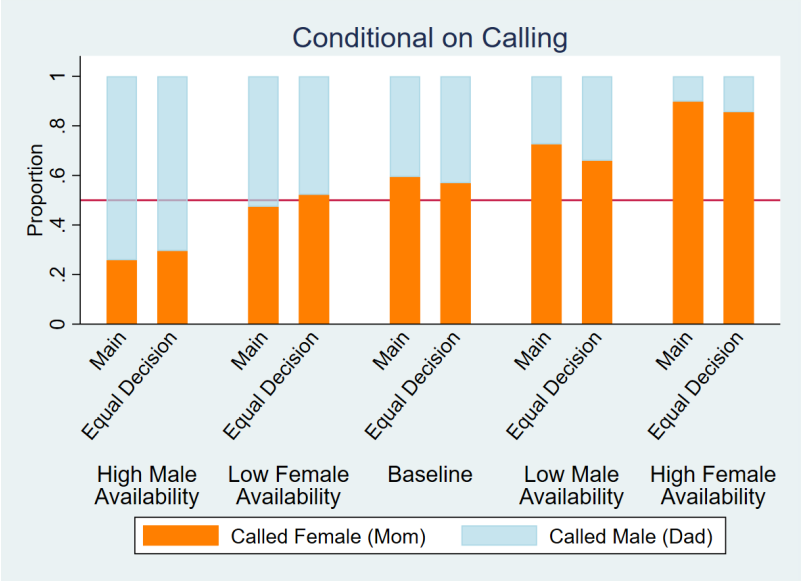
	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
Elementary	0.50	0.50	0.50	0.49	0.52
Middle	0.15	0.14	0.16	0.15	0.17
High	0.19	0.19	0.22	0.21	0.20
Decison-Maker Female	0.58	0.60	0.58	0.58	0.58
PublicCharter	0.06	0.07	0.05	0.06	0.06
PublicNOTCharter	0.78	0.75	0.81	0.78	0.81
Private	0.17	0.18	0.14	0.16	0.12
FreeLunch	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: 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.

Figure J4: Outcomes By Treatment "Main" vs. "Equal Decision" Variations



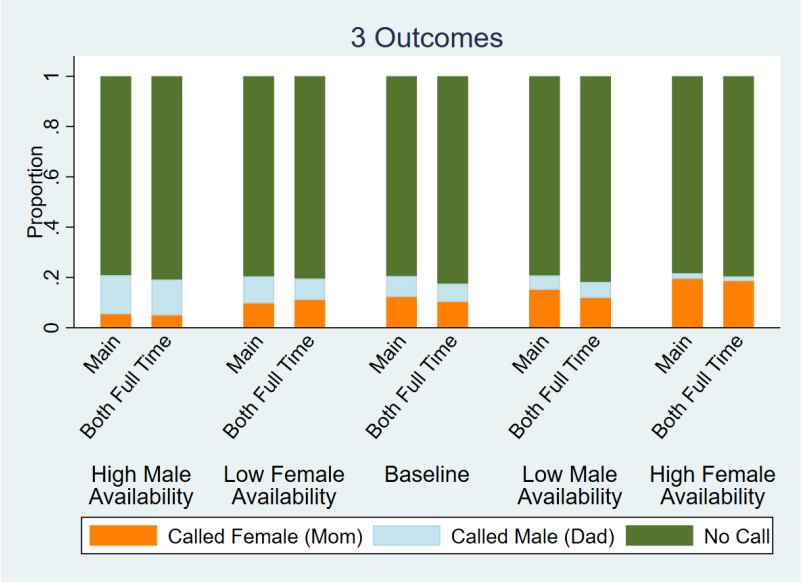
(a) All Outcomes



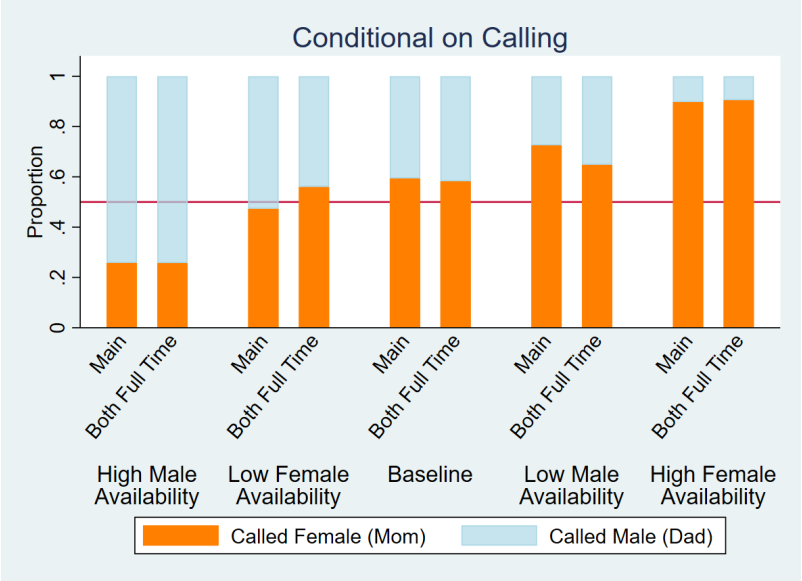
(b) Outcomes Conditional On Calling

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 "This is the type of decision we both want to be involved in equally." 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 Main Variation. Panel (a) represents three outcomes from $30,471$ decision-makers in Main and $30,320$ in Equal Decision, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 7,778$ in Main and $7,209$ in Equal Decision). If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

Figure J5: Outcomes By Treatment "Main" vs. "Full Time" Variations



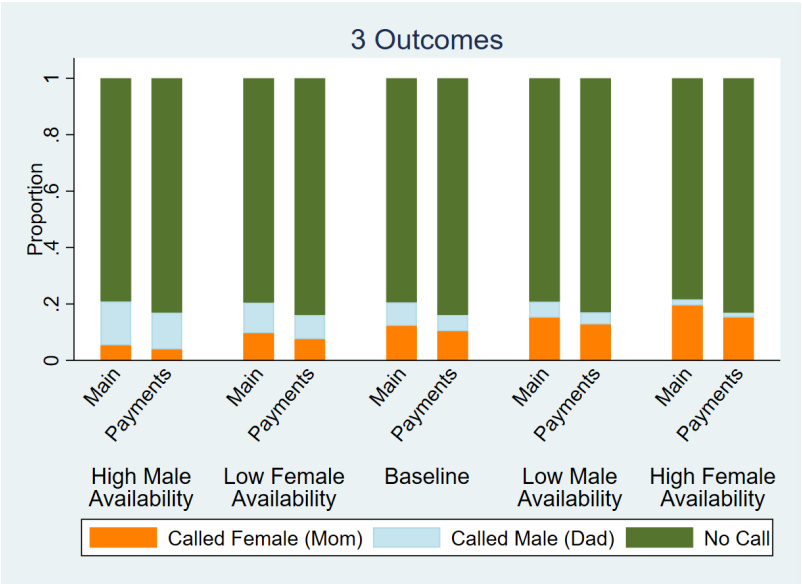
(a) All Outcomes



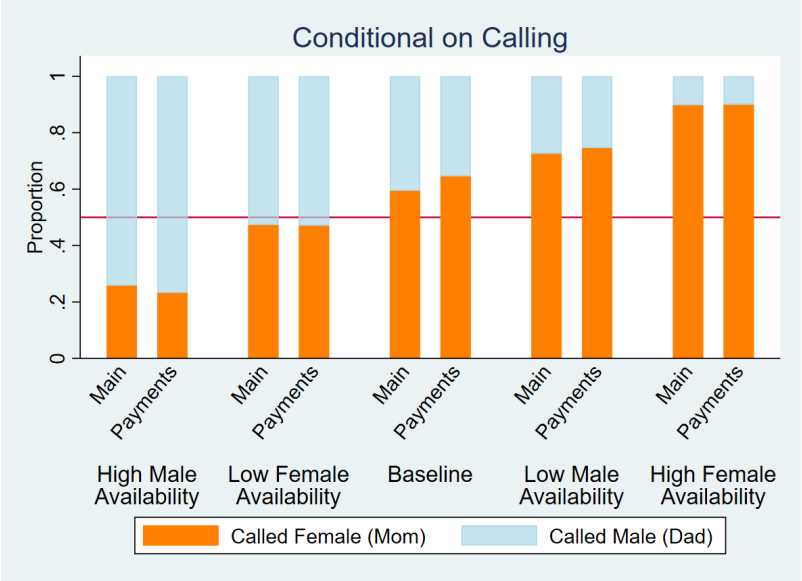
(b) Outcomes Conditional On Calling

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 Main Variation. Panel (a) represents three outcomes from $30,471$ decision-makers in Main 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 ($N = 7,778$ in Main and $2,175$ in Full Time). If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

Figure J6: Outcomes By Treatment "Main" 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 Main Variation. Panel (a) represents three outcomes from $30,471$ decision-makers in Main 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 ($N = 7,778$ in Main and $9,472$ in Full Time). If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

K Addressing Imbalance In Emails Sent From Mothers and Fathers

To address the unintentional imbalance in emails sent from mother's emails versus father's email in the main text we have weighted our observations so that emails from each parent are balanced. In this section we randomly delete observations from our data until we have achieved balance on emails from each parent, and find very similar results to those reported in the main text.

Table K8: **Summary Statistics By Treatment in Main Variation Similar to Table 1**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
FemaleNum0	0.05	0.10	0.12	0.15	0.19
MaleNum0	0.16	0.11	0.08	0.06	0.02
NoCall	0.79	0.79	0.80	0.79	0.78
FemaleNum	0.26	0.47	0.59	0.73	0.90
MaleNum	0.74	0.53	0.41	0.27	0.10
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	6728	5486	5016	5618	6066

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother.

Table K9: **Summary Statistics By Variation (All Treatments Combined) Similar to Table 2**

	(1) Main	(2) Equal Decision	(3) Full Time	(4) Payments
FemaleNum0	0.122	0.121	0.117	0.100
MaleNum0	0.087	0.080	0.074	0.068
NoCall	0.791	0.799	0.808	0.832
FemaleNum	0.582	0.602	0.613	0.596
MaleNum	0.418	0.398	0.387	0.404
FemaleEmail	0.500	0.501	0.499	0.499
Observations	28914	28692	7983	8443

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother.

Table K10: **Summary Statistics By Primary Email Sender Similar to Table 3**

Panel A: Email Sent By Mother (CCing Father)

	(1) All Messages	(2) High Male (Hm)	(3) Low Female (Lf)	(4) Baseline (b)	(5) Low Male (Lm)	(6) High Female (Hf)
FemaleNum0	0.17	0.08	0.18	0.20	0.21	0.20
MaleNum0	0.04	0.13	0.03	0.00	0.01	0.01
NoCall	0.79	0.78	0.79	0.80	0.78	0.79
FemaleNum	0.81	0.38	0.86	0.98	0.96	0.97
MaleNum	0.19	0.62	0.14	0.02	0.04	0.03
FemaleEmail	1.00	1.00	1.00	1.00	1.00	1.00
Observations	14448	3365	2726	2512	2813	3032

Panel B: Email Sent By Father (CCing Mother)

	(1) All Messages	(2) High Male (Hm)	(3) Low Female (Lf)	(4) Baseline (b)	(5) Low Male (Lm)	(6) High Female (Hf)
FemaleNum0	0.07	0.02	0.02	0.04	0.10	0.18
MaleNum0	0.13	0.18	0.19	0.16	0.11	0.04
NoCall	0.79	0.80	0.79	0.80	0.80	0.78
FemaleNum	0.35	0.12	0.08	0.21	0.48	0.83
MaleNum	0.65	0.88	0.92	0.79	0.52	0.17
FemaleEmail	0.00	0.00	0.00	0.00	0.00	0.00
Observations	14466	3363	2760	2504	2805	3034

Notes: FemaleNum0 is the proportion of calls made to a female parent when including NoCall as an outcome, and MaleNum0 is defined analogously. FemaleNum is a variable that takes a null value if no call was made, and thus represents the proportion of calls to female parents conditional on a call being made; MaleNum is defined analogously. FemaleMaleRatio can be computed either as the ratio of FemaleNum0 to MaleNum0 or as FemaleNum to MaleNum. FemaleEmail takes the value 1 if a the email was sent from the mother's email address and CCed the father, and the value 0 if the email was sent from the father and CCed the mother.

Table K11: **Multinomial Logit Models of Effect of Treatments on No Call, Call Male or Call Female Similar to Table A1**

	(1)	(2)	(3)	(4)	(5)	(6)
	outcome	outcome	outcome	outcome	outcome	outcome
No_Call						
High Male (Hm)	-0.63***	-0.67***	0.80***	0.83***	0.00	0.00
	(0.06)	(0.06)	(0.07)	(0.07)	(.)	(.)
Low Female (Lf)	-0.28***	-0.29***	0.21**	0.22***	0.00	0.00
	(0.07)	(0.07)	(0.06)	(0.06)	(.)	(.)
Low Male (Lm)	0.37***	0.38***	-0.25***	-0.26***	0.00	0.00
	(0.08)	(0.08)	(0.06)	(0.06)	(.)	(.)
High Female (Hf)	1.30***	1.34***	-0.50***	-0.53***	0.00	0.00
	(0.10)	(0.10)	(0.05)	(0.06)	(.)	(.)
Female_Call						
High Male (Hm)	-1.43***	-1.50***	0.00	0.00	-0.80***	-0.83***
	(0.09)	(0.09)	(.)	(.)	(0.07)	(0.07)
Low Female (Lf)	-0.49***	-0.51***	0.00	0.00	-0.21**	-0.22***
	(0.09)	(0.09)	(.)	(.)	(0.06)	(0.06)
Low Male (Lm)	0.61***	0.64***	0.00	0.00	0.25***	0.26***
	(0.09)	(0.09)	(.)	(.)	(0.06)	(0.06)
High Female (Hf)	1.80***	1.87***	0.00	0.00	0.50***	0.53***
	(0.11)	(0.11)	(.)	(.)	(0.05)	(0.06)
Male_Call						
High Male (Hm)	0.00	0.00	1.43***	1.50***	0.63***	0.67***
	(.)	(.)	(0.09)	(0.09)	(0.06)	(0.06)
Low Female (Lf)	0.00	0.00	0.49***	0.51***	0.28***	0.29***
	(.)	(.)	(0.09)	(0.09)	(0.07)	(0.07)
Low Male (Lm)	0.00	0.00	-0.61***	-0.64***	-0.37***	-0.38***
	(.)	(.)	(0.09)	(0.09)	(0.08)	(0.08)
High Female (Hf)	0.00	0.00	-1.80***	-1.87***	-1.30***	-1.34***
	(.)	(.)	(0.11)	(0.11)	(0.10)	(0.10)
Control Variables						
R^2						
Observations	28914	28914	28914	28914	28914	28914

Notes: This table presents the results of a multinomial logit model using a model like the one in Equation 10. 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 in columns (1) and (2), female call in columns (3) and (4), and male call in columns (5) and (6). The results from the three base cases are analogous and all three are presented to make specific comparisons more simple. The outcomes with no controls from this table are represented visually in Figure 4.

Table K12: **Multinomial Logit Models For Theory Model Similar to Table A2**

	(1) Main	(2) Equal Decision	(3) Full Time	(4) Payments
No_Call				
any_msg_M	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
x_M	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
any_msg_F	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
x_F	0.00	0.00	0.00	0.00
	(.)	(.)	(.)	(.)
Female_Call				
any_msg_M	-0.28***	-0.17**	-0.28*	-0.34**
	(0.05)	(0.05)	(0.11)	(0.11)
x_M	-0.52***	-0.37***	-0.44***	-0.58***
	(0.03)	(0.04)	(0.07)	(0.07)
any_msg_F	0.15**	0.33***	0.36***	0.06
	(0.05)	(0.05)	(0.10)	(0.10)
x_F	0.35***	0.35***	0.28***	0.36***
	(0.03)	(0.03)	(0.05)	(0.06)
Male_Call				
any_msg_M	0.13*	0.19***	0.33**	0.36**
	(0.06)	(0.06)	(0.12)	(0.13)
x_M	0.50***	0.40***	0.40***	0.51***
	(0.03)	(0.03)	(0.06)	(0.06)
any_msg_F	-0.51***	-0.34***	-0.52***	-0.34*
	(0.07)	(0.06)	(0.14)	(0.15)
x_F	-0.79***	-0.49***	-0.75***	-0.81***
	(0.05)	(0.04)	(0.10)	(0.11)
Control Variables				
R^2				
Observations	28914	28692	7983	8443

Notes: This table presents the results of a multinomial logit model using a model like the one in Equation 10. 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. The right hand side variables are discussed in Section 2.

Table K13: **Balance on Observable Attributes of Schools/Decision Makers By Treatment In Main Variation Similar to Table D3**

	(1) High Male (Hm)	(2) Low Female (Lf)	(3) Baseline (b)	(4) Low Male (Lm)	(5) High Female (Hf)
Elementary	0.48	0.49	0.51	0.50	0.50
Middle	0.14	0.14	0.14	0.15	0.15
High	0.19	0.20	0.20	0.19	0.20
Decison-Maker Female	0.57	0.58	0.59	0.59	0.58
PublicCharter	0.06	0.05	0.06	0.06	0.06
PublicNOTCharter	0.76	0.79	0.81	0.79	0.80
Private	0.18	0.16	0.13	0.15	0.14
FreeLunch	0.55	0.56	0.54	0.55	0.53
White	0.53	0.52	0.52	0.53	0.52
Black	0.14	0.15	0.14	0.14	0.15
Hispanic	0.23	0.22	0.23	0.23	0.23
FemaleEmail	0.50	0.50	0.50	0.50	0.50
Observations	6728	5486	5016	5618	6066

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.