Who You Gonna Call?
Gender Inequality in External Demands for Parental Involvement*

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

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

July 10, 2024

Abstract

Gender imbalance in time spent on child rearing causes gender inequalities in labor market outcomes, human capital accumulation, and economic mobility. We investigate a novel source of this inequality: external demands for parental involvement. We pair a theoretical model with a large-scale field experiment with a near-universe of US schools. Schools receive an email from a two-parent household with a general inquiry and are asked to call one of the parents back. Mothers are 1.4 times more likely than fathers to be contacted. We decompose this inequality into discrimination stemming from differential beliefs about parents' responsiveness versus other factors, including gender norms and link it to the gender earnings gap and other labor market outcomes. Our findings underscore a process through which agents outside the household contribute to within-household gender inequalities.

JEL Classification: J16, J71, C93, J22
Keywords: Discrimination, Gender Gap, Field Experiment

*We thank Katelyn Cranney, Mason Fields, Katherine Jolley, Boston Kelley, Paris Rich, and Kiersten Robertson for their excellent research assistance. We are grateful to Amanda Agan, Nageeb Ali, Joshua Angrist, David Autor, Peter Bergman, Fran Blau, Kasey Buckles, Amanda Chuan, Zoe Cullen, Amy Finkelstein, Ingrid Haegele, Henning Hermes, Christopher Karpowitz, Lars Lefgren, Louis-Pierre Lepage, Fabian Mierisch, Stephen O'Connell, Emily Oster, Daniele Paserman, Amanda Pallais, Jessica Preece, Heather Royer, Nina Roussille, Frank Schilbach, and Betsey Stevenson for their valuable feedback. 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, Behavioral Experimental Design Initiative, Stockholm University (SOFI), Uppsala University, Gothenburg University, Middlebury College, MIT, Montana State University, UMass Amherst, Georgia State University, University of Delaware, Loyola Marymount University, University of California at Santa Barbara, University of Colorado Denver, George Mason University, IPPF Congress, Nebraska Labor Summit, and the Association for Social Sciences (ASSA) meetings. We thank the Brigham Young University College of Family, Home, and Social Sciences, the Syracuse University Maxwell School, and Tufts University for their generous research support. This work has been supported (in part) by Grant 2204-38170 from the Russell Sage Foundation, and the Institute for Human Studies IHS017750. Any opinions expressed are those of the principal investigators alone and should not be construed as representing the opinions of the funders. This study was pre-registered at the AEA registry (AEARCTR-0007610) and was approved by the relevant IRBs. A previous version of this paper was circulated under the title "Who Ya Gonna Call? Gender Inequality in External Demands for Parental Involvement."

1 Introduction

Despite the convergence of men's and women's roles in the labor market, a substantial and persistent gender earnings gap of nearly 18% remains (US Census Bureau, 2020). Prior studies have documented many factors contributing to this gap. Of recent focus is women's tendency to concentrate in occupations with more temporal flexibility, which is especially true for women with children (Price and Wasserman, 2022; Duchini and Van Effenterre, 2022; Wasserman, 2022; Goldin, 2014).

The need for greater workplace flexibility is consistent with the robust finding that women—even those who work outside the home—engage in a disproportionate share of child- and household-related tasks.[1] US Time Use Survey data reveal that married mothers employed full time spend over 50% more time caring for children and engaging in housework and food preparation than analogous fathers (see panel (a) of Figure 1). Similarly, Cubas et al. (2021) find that 35% of mothers experience a household interruption during their workday, compared to only 20% of fathers. These gender imbalances have significant economic costs to women, stunting labor market outcomes, human capital accumulation, and economic growth, as documented extensively in the motherhood wage gap literature.[2]

In this paper, we investigate a novel source of this inequality, which we refer to as "external demands for parental involvement." In short, institutions beyond the household and beyond the place of employment put demands on families, and these demands may fall disproportionately on mothers. These external demands come from outside forces, such as schools, doctors' offices, extracurricular activities, or even grandparents. Small, optimizing decisions by these external agents could create powerful disadvantages for women as they anticipate and respond to these external demands by changing the type of work they do, the careers they choose, and how they progress in their careers. These biases also limit men's ability to be fully involved in their children's lives and disadvantage a growing number of fathers who report wanting a more equal distribution of these external demands.[3] This might lead to worse outcomes for society as a whole by reinforcing social biases and perpetuating the cycle of gender inequality.

The social biases inherent in the external demands placed on parents can take many forms.

[1] See, for example, Aguiar and Hurst (2007); Craig and Mullan (2011); Schoonbroodt (2018).

[2] Many prior studies have documented the motherhood wage gap in a wide range of contexts, including work by Adams-Prassl et al. (2023); Ciasullo and Uccioli (2023); Kleven (2023); Speer and Ersoy (2022); Jack et al. (n.d.); Erosa et al. (2022); Albanese et al. (2022); Cubas et al. (2021); Duchini and Van Effenterre (2022); Cubas et al. (2022); Kleven et al. (2019); Kuziemko et al. (2018) and Angelov et al. (2016).

[3] A recent nationally representative survey finds that men want to be contacted by their child's school 47% of the time (American Family Survey, 2022).

For example, women may get called on more often than men for child-related tasks, such as school-related requests. Schools, therefore, provide an ideal setting to investigate external demands for parental involvement by gender. To do this, we develop a theoretical model which informs the design of a field experiment in a K-12 school setting. Specifically, we send emails with phone numbers for both parents in a fictitious two-parent household to the near-universe of US school principals (N = 80,071), asking the principal to contact a parent by phone about a general school-related inquiry. We randomly vary which parent sends the email and the information about their availability to disentangle whether discrimination stems from decision-makers' beliefs about parents' responsiveness or other deterrents. Beliefs about responsiveness might include the perception that women are more available because they are stay-at-home mothers or that women naturally want to be more involved in a school-related decision and will, therefore, be more responsive than men. Other deterrents might include distaste for calling a specific parent, systemic factors, social norms, or beliefs not related to the value of a parent's response.

Experimentally varying the information about parents' availability and desire for equal decision making allows us to investigate whether the gender gap can be mitigated by households changing the signals they send. As our experiment shows, signaling parents' responsiveness goes only so far in affecting the desired change. Our model allows us to further explore other attributes at play, such as the prevailing gender norms of schools and geographic locations. We show that such attributes impact the inequality in demands on parents' time, implying the gender gap might be mitigated through policies targeting behavioral change in specific sub-groups.[4]

We find striking gender and treatment differences. Principals are significantly more likely to call mothers first in our simplest message, which contains no information about parents' availability. On average, conditional on a call being made, mothers are called first 1.4 times more than fathers (59% versus 41% for the 20% of principals who make any call), providing direct and novel evidence of greater external demands on mothers relative to fathers. Our findings underscore a significant gender inequality in external demands, which are plentiful within the school setting (e.g., picking up a sick child, volunteering for school events) and beyond (e.g., which parent schedules doctor visits and registers for summer camps, who coordinates extracurricular activities, and who grandparents expect to take care of a child's needs) indicating a significant impact on mothers' time and labor market outcomes. Using

[4]The scope of this paper is exclusive to two-parent households with a male and female parent. We acknowledge that there are many types of households and more gender identities, but we believe that work using the two extreme ends of the gender spectrum (male/female) is an important first step in exploring how gender identity affects external demands on a person's time. Furthermore, we believe that exploring the effect of external demands in other settings is an important question for future work, and we discuss this in Appendix H.

a survey of parents as well as ATUS data, we link the inequality in external demands to the gender earnings gap and show that women are more likely than men to incur career penalties as a result. We also document that even when households exert substantial efforts to achieve a more balanced split of child-related tasks (eg. by repeatedly reminding the school who to contact or by outsourcing the task), they incur disruption costs which tend to exacerbate existing gender gaps in the labor market.

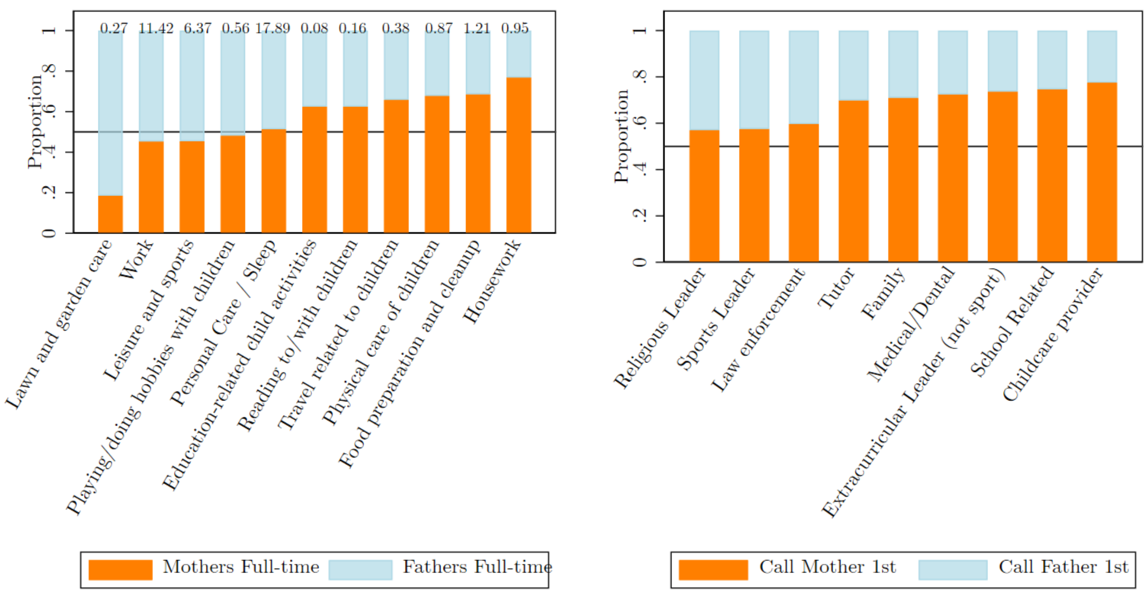
Finally, in addition to documenting a significant gender gap in external demands for parents' time, we explore the reasons why it arises and test potential mechanisms. Specifically, we show that signaling that the father is more available mitigates the inequality and causes mothers to be called less than half the time. It is notable, however, that even when fathers signal that they are more available, mothers still get 26% of the calls. In contrast, signals that reinforce stereotypes that mothers are more available cause them to receive 90% of the calls. Sending the email from the father significantly raises the share of calls to fathers. However, even when the email comes from the father and contains a positive signal about father's availability, 12% 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.

To identify the mechanisms underlying any differential demand for parental involvement, we pair a novel theoretical model with our field experiment as well as multiple surveys of parents and individuals whose jobs involve interacting with parents and children. Our model shows how decision-makers choose whether to contact a mother, father, or neither parent. It allows us to attribute any differences we find to statistical discrimination on the basis of beliefs about parents' responsiveness, norms about calling the person who made contact first, or other factors, which we identify through a separate survey with school administrators. Our randomized signals about parents' availability and desire for equal decision-making only impact a decision-maker's beliefs about the benefit of a call due to changes in responsiveness. Thus, the differences in the proportions of calls across signals tell us what happens when those beliefs change. Any residual differences in the proportions of calls to mothers versus fathers are attributable to other deterrents.[5] We find that differential beliefs about parents' responsiveness and differences in other deterrents account for roughly equal shares of the excess calls to mothers.

This paper extends the existing literature in four important ways. First, we experimentally document a novel gender gap: that is, gendered differences in external demands for parental involvement. While prior research has found that women spend significantly more time on

[5]Our model also accounts for the impact of which parent sends the email, but this impact is neutralized in our results because our intervention is balanced with respect to which parent sends the email.

Figure 1: Gender Inequality in Household Time Use and External Contacts



(a) Proportion of Time Full-time Employed Mothers vs. Fathers in Two Parent Households Spent in Day (48 Hours Per Household)

(b) Proportion of Time Mother vs. Father in Two Parent Households are Contacted First By Type of External Decision-Maker

Notes:
Panel (a) shows the proportion of time spent by male versus female respondents on different activities. Respondents are married adults working full-time with children under 18 from the American Time Use Survey from the BLS years 2015-19 combined. There is a line at the equal time spent on an activity by male versus female respondents. The number at the top of each bar is the total hours spent on this activity by male and female respondents collectively (sums close to 48 hours). For brevity, we exclude some categories (e.g., purchasing goods/services, caring for non-children, non-child related travel, and other activities). Full time working mothers tend to spend equal or more time on these excluded categories relative to the full-time working fathers.
Panel (b) shows the proportion of time mothers and fathers are contacted by adult leaders who interact with parents. There is a line at the equal amounts of contact to mothers versus fathers. Respondents were 300 adults who interacted with parents and self-identified as doing so mostly within a certain role (e.g., Teacher, Nurse, Sports Leader). See Appendix L.1 for details. We told respondents to imagine "a family that consists of one mother and one father living together jointly raising at least one child." We then asked respondents the following question about a mother or a father: What proportion of the time do you contact the [father][mother] first if only contacting one parent first? with 50% being randomized to be asked about the [father], and 50% randomized to be asked about the [mother].

child-related tasks than men in two parent households, our study is the first to show that this inequality is, in part, driven by external demands for parental involvement. This gender inequality has substantial economic and social costs for women and men, who both report a desire for a more equal distribution of child-related tasks (Pew Research Center, 2015). A nationally representative survey of parents of school-age children finds that women report being contacted by the school more often than men, yet wish they were contacted less often, while men wish to be contacted about half the time (American Family Survey, 2022). We also find that women are significantly more likely to be the point of contact for external decision-makers across a wide range of child-related domains, from doctors' offices to extracurricular sports coaches to religious leaders (see panel (b) of Figure 1).[6] Perhaps most importantly, in our own survey (Appendix L.3), mothers were significantly more likely than fathers to report that child-related external interruptions negatively impacted their careers and earnings.

Related prior research has documented the effects of childcare and other care-giving disruptions on women's labor market outcomes. Price and Wasserman (2022), for example, show that summer childcare constraints contribute to career choices and earnings for women with school-aged children, in line with findings from Duchini and Van Effenterre (2022) and Cowan et al. (2023). Similarly, the COVID-19 pandemic and the associated school and daycare closures 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] 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.

Second, we contribute to the growing literature on the role of information 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 that providing accurate information reduces statistical discrimination (Laouénan and Rathelot, 2022; Bohren et al., 2019), others have found no discernible effects (Bertrand and Mullainathan, 2004; Oreopoulos, 2011). Our paper advances this literature by documenting a notable 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 the availability of fathers moves calls away from mothers, we also document the limits of this informational intervention. Specifically, in our baseline variation, we find that signaling mothers' high availability leads to mothers being contacted 90% of the time, while signaling fathers' high availability increases calls to fathers only up to 74%.

[6]Prior studies suggest that women anticipate greater external demands for parental involvement long before having children which may push them toward more flexible jobs, leading to substantial labor market penalties, including reduced labor force participation (Kleven et al., 2021; Mas and Pallais, 2020; Bursztyn et al., 2017; Mas and Pallais, 2017; Pertold-Gebicka et al., 2016; Anderson et al., 2002) and curbed earnings (Cortes and Pan, 2021; Goldin, 2014; Gicheva, 2013).

[7]Adams-Prassl et al. (2023); Couch et al. (2022); Garcia and Cowan (2022); Hansen et al. (2022); Amuedo-Dorantes et al. (2020); Zamarro and Prados (2021); Sevilla and Smith (2020); Montes et al. (2021); Heggeness (2020); Russell and Sun (2020); APA (2021).

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 experiments can 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; Bohren et al., forthcoming) with recent work emphasizing the importance of indirect discrimination stemming from systemic and institutional factors (Bohren et al., 2022). We join a small but growing literature that attempts to differentiate these sources of discriminatory behavior.

Prior research has employed field experiments to tease out the true sources of discriminatory behavior. For example, List (2004) examines racial discrimination in the baseball card market, Islam et al. (2018) investigate how patients choose a physician, and Bohren et al. (2019) examine gender discrimination in a mathematics forum. We advance this literature by using a simple, static theoretical model combined with a field experiment to identify separate parameters that capture the availability beliefs versus other deterrents that lead to discriminatory behavior. While we cannot distinguish between accurate and inaccurate beliefs in our setting, using the structural model to identify beliefs allows us to avoid the identification problem present in many studies attempting to isolate the source of discrimination (Bohren et al., forthcoming).

Finally, this paper contributes to the literature on institutional, 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; Karpowitz et al., 2023; Kline et al., 2022; Babcock et al., 2017; Scott, 2013; Council, 2004; Powell and DiMaggio, 2012). We provide novel evidence of systemic discrimination by showing that school principals' optimizing behavior ends up creating worse outcomes for some individuals in society and arguably for society as a whole. As Small and Pager (2020) argue, institutional discrimination deserves particular attention, given the deeply ingrained nature of systemic practices and their long-lasting consequences.

2 Field Experiment

Our theoretical model (discussed in Section 4) and a survey of educators inform the design of a large-scale field experiment, which consists of sending email messages to a near-universe of US school principals. The emails are sent from a set of fictitious parents, one male and one female.[8] Email is a common way for parents to contact schools; in our survey, 75% of educators report being contacted by parents via email at least once a month.[9] Our specific inquiry is meant to mimic the type of message a household might send when relocating to a new area and exploring new school options. 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 (see, for example, Diaz-Serrano and Meix-Llop, 2016; Bergman and McFarlin Jr, 2018; Ahmed et al., 2020; Oberfield and Incantalupo, 2021; Cantet et al., 2022; and Hermes et al., 2023).

In the study most like our own, Hermes et al. (2023) email childcare centers in Germany from either the mother or the father and find that response rates are similar, but responses to mothers are shorter and less positive than responses to fathers. Importantly Hermes et al. (2023) do not offer decision-makers the choice between contacting a mother or a father, like we do in the current study, so our outcome variables are not directly comparable. Furthermore, they explore responses from parents about optional childcare for young children while we look at questions about mandatory schooling for older children. Arguably, early childcare is a setting where a woman contacting a childcare center might be viewed as shirking her maternal responsibilities, while in our setting, a woman contacting a school might be viewed as an involved parent. Consistent with our results, they find that gender norms may be a major driver of the observed inequality. We discuss this further in Section 5.1.

2.1 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, and 97% of parents send their children to school outside the home (NCES, 2021). Schools are an ideal setting to explore external demands on parents' time because of its' near-universal relevance and because 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).

For several reasons, we believe that any gender gaps that we document in our specific task 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 (PTA). First, educators in our survey say that they would contact the mother first in many of these scenarios (we discuss

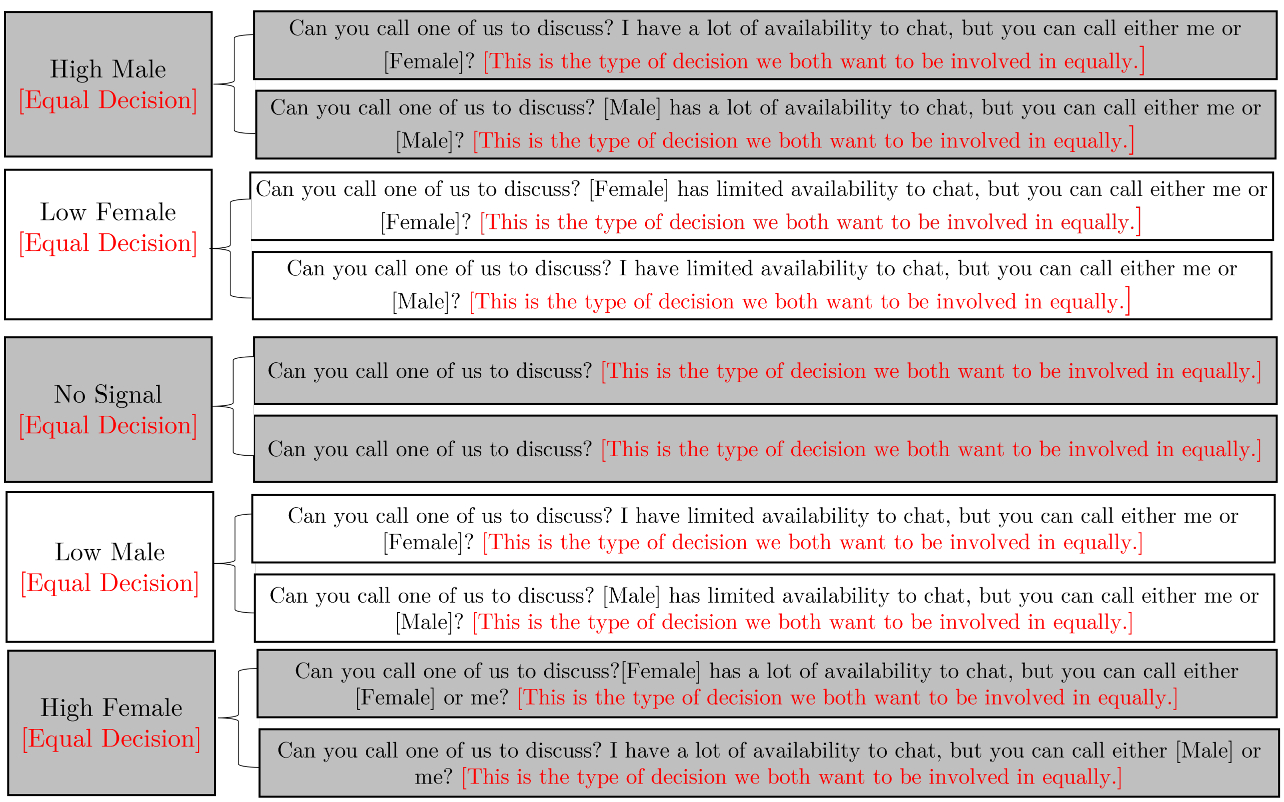
[8]We describe our data collection process in more detail in Appendix K as well as some of the ethical considerations in Appendix J.

[9]We discuss the survey in detail in Appendix L.1.

the survey in Appendix L.1). Second, the gender distribution of these tasks is significantly skewed; mothers comprise almost 90% of PTA members, and only 13% of fathers report high levels of involvement in their child's school activities, compared to 53% of mothers.[10]

Furthermore, although the gender gap in external demands for parental involvement is established for our test case of outreach from a school administrator, we expect that it indicates a dynamic likely present in a wide range of social situations that require parental attention or input. As shown in Figure 1, mothers spend more time than fathers on many tasks, and decision-makers from various organizations beyond schools report contacting mothers more than fathers.

Figure 2: Field Experiment Variation in Messages



Notes: In this figure, we show a pertinent portion of differences in the messages we sent to schools in both our Baseline and Equal Decision variations. Within each variation, there are five treatment messages: High Male, Low Female, No Signal, Low Male, and High Female. The parent who sent the email always had their phone number listed first. Above, we show the message sent from the male parent (cc'ing the female parent) and then the message from the female parent (cc'ing the male parent). The full text of example email messages in the Baseline variation is available in Appendix Section F.

[10] See Daly and Groes (2017); Belkin (2009); Scottland (2020).

2.2 Messages

In our experiment, school principals receive an email from a fictitious 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. The email sender's phone number is always listed first, and we randomize whether the primary sender is the father or mother. We call this the "No Signal" message.[11] We developed the specific message in consultation with school administrators from various schools (public, private, and charter). Our conversations and survey evidence (Appendix L.1) confirmed parents frequently make general email inquiries to schools before enrolling and that it is common for one parent to email, copying the other parent.

We then augment our No Signal message in two ways. First, we add a baseline sentence indicating the availability of a specific parent in the two-parent household. Figure 2 shows the exact variation in wording. Details of the exact names and email addresses used in the experiment are in Appendix K, and the full text of the Baseline variation messages is in Appendix F.[12] This leaves us with a total of five Baseline messages.

Second, because messages about availability might also send signals about a desire for equal decision-making, we send an additional five messages that add on a sentence meant to fix beliefs about the household's preferences for equal decision-making. Specifically, we add the following sentence, "This is the type of decision we both want to be involved in equally." In Appendix H, we discuss variations of these messages (e.g., longer and more detailed) that we sent to a sub-sample of principals. Our findings are robust to these variations.

We designed these messages based on our theoretical model discussed in Section 4 as well as a survey we conducted with educators detailed in Appendix L.1. Our survey findings reveal that a key dimension on which educators could be statistically discriminating is differential beliefs about mothers' relative responsiveness. Specifically, common reasons educators gave for calling mothers first were, "I expect this person to be more likely to respond quickly" and "This person is more interested/willing." One of the model's key results is that, by varying the strength (low/high) of the signals about each of our parents' availability, as well as their desire for equal decision-making, we can disentangle the extent to which

[11] To be precise, it is a "No Verbal Signal" message, and there is a non-verbal signal inherent in which parent sends the email. We will address this issue later.

[12] One might wonder specifically about the realism of messages which are from Parent A but then state that Parent A is not very available (e.g. Low Female sent from mother or Low Male sent from father). However, as seen in Table 2, response rates for these emails are quite similar to the other emails sent, which seems in line with no difference in realism for these email messages versus our other messages.

the gender inequality is driven by differential beliefs about parents' responsiveness.[13]

Our emails also contain a key nonverbal signal: which parent sends the email. Many survey respondents stated that they would call the parent who is listed first or who reaches out to them. Our analysis will, therefore, consider which parent sends the email, and we will allow this effect to vary by treatment and variation since decision-makers' willingness to override this rule of thumb may vary depending on the verbal content of the message they receive.

2.3 Sample Frames and Data Collection

During the summer of 2022,[14] we sent emails to a near-universe (a sample of 80,071) of school principals across the US. We begin by describing the Baseline and Equal Decision variations of our experiment, which were sent to over 60,000 school principals. 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, content, and length of any voicemail 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 K 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.

Our primary outcome of interest is whether a decision-maker calls the female parent, the male parent, or neither parent. Decision-makers can also email or text our parents; however, we set up an auto-response to emails and texts and found that fewer than 0.2% of our principals responded via a text message. 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:

[13]In a previous draft of this paper, we presented a preliminary version of our theoretical model that did not account for the effect of which parent sends the email. We have added this to the model in response to feedback. We now also focus on the version of the model with messages that fix beliefs about preferences for equal decision-making.

[14]Throughout 2021, we conducted a series of pilot experiments with a total of 3,267 observations to iron out implementation logistics. Some pilot emails were sent out during the school year, while others were sent during the summer. Notably, we did not observe significant differences in response rates by time of year.

$$p_{ij}(x) = \frac{e^{\beta_j^{l\boldsymbol{M}}(LowMale)+\beta_j^{h\boldsymbol{M}}(HighMale)+\beta_j^{l\boldsymbol{F}}(LowFemale)+\beta_j^{h\boldsymbol{F}}(HighFemale)+\boldsymbol{\alpha}X_i}}{\sum_{k\in n,f,m} e^{\beta_k^{l\boldsymbol{M}}(LowMale)+\beta_k^{h\boldsymbol{M}}(HighMale)+\beta_k^{l\boldsymbol{F}}(LowFemale)+\beta_k^{h\boldsymbol{F}}(HighFemale)+\boldsymbol{\alpha}X_i}}. \tag{1}$$

In this regression model, p_{ij} is the probability that individual i calls neither parent ($j = n$), the female parent ($j = f$), or the male parent ($j = m$). Next, we have treatment indicators for each treatment beyond the No Signal treatment: LowMale, HighMale, LowFemale, and HighFemale. We can also include a vector X_i of covariates, including which parent sent the email (cc'ing the other parent) and attributes of the decision-maker and their school.

In subsequent analysis, we let the outcome variable instead be binary, taking the value one when a female parent is called and zero otherwise. We then run a simple linear regression for ease of interpreting the coefficients.

3 Results: Gender Inequality & Signal Impact

We are balanced on observable variables across our treatments as shown in Tables C.1 and C.2. Although we had intended to send an equal number of emails from fathers and mothers, as well as an equal number of emails in each of our treatments, these design choices were not attained due to some computing errors.[15] Our results are based on re-weighted data such that there is balance in the number of messages sent in each of our five messages (Figure 2), and there is balance between the number of messages sent from fathers versus mothers within a treatment arm. However, our results are quantitatively and qualitatively the same when we randomly exclude observations to achieve balance (available from authors upon request)

Overall, we observe a 20% response rate from the principals, which is in line with previous work. Recent studies where job applicants submit applications with a phone number and email to employers find response rates from employers of about 8% to 11% (Agan and Starr, 2018). In studies with 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 70% (Diaz-Serrano and Meix-Llop, 2016; Bergman and McFarlin Jr, 2018; Ahmed et al., 2020; Oberfield and Incantalupo, 2021; Cantet et al., 2022;

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

Hermes et al., 2023). Another related outcome is whether principals take a survey in response to an email request, where recent work finds only 14% of principals take this action (Neal et al., 2020).

We compare the observable characteristics of the principals who call back with those who do not and find mostly small but statistically significant differences. As reported in Appendix Table A.2, we are less likely to get a call back from elementary schools, public (excluding charter) schools, and schools with a higher share of students receiving free lunch and non-white students. While this suggests selection into calling, we believe much of this selection is driven by the relatively fewer resources at public schools versus private schools and schools with a higher share of non-white students, as well as students receiving free lunch. If anything, we believe that elementary schools are more likely to call mothers than other types of schools, which likely biases us toward measuring too little of a mother preference in our data.

3.1 Gender Inequality with No Signal

Table 1 and Figure 3 and report the proportion of actions taken by decision-makers in all of our conditions, including the No Signal conditions (column (3) of Table 1 or center bars of Figure 3), which contain no verbal information about parents' availability. If there was no gender inequality and decision-makers were randomly choosing which parent to call, we would expect the same proportion of calls to male and female parents. In our No Signal Baseline variation message, we observe that about 12% of school principals call mothers first, while only 8% call fathers first. The remaining decision-makers do not call either parent. The difference in calls to male and female parents is large and statistically significant ($Pr(T > t) = 0.00$). 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 toward calling female parents is in the ratio of female-to-male calls in the No Signal messages, which is about 1.4. This is well above the ratio of 1 that we would expect if decision-makers were randomizing which parent to call, and it means that mothers are 1.4 times more likely than fathers to receive a call. Conditional on receiving a callback, mothers are called first about 60% of the time in No Signal treatment (both with and without the addition of the sentence about making decisions equally).

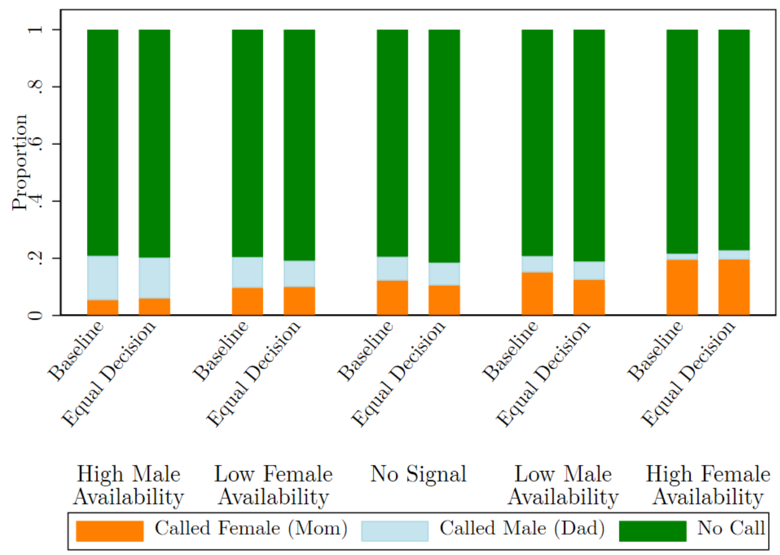
The gender gap we document is almost surely a lower bound estimate of the gender inequality in external demands from schools for several reasons. First, our experiment essentially sends an equal number of requests from mothers and fathers, neutralizing any gender

Table 1: **Summary Statistics by Treatment in Baseline & Equal Decision Variation**

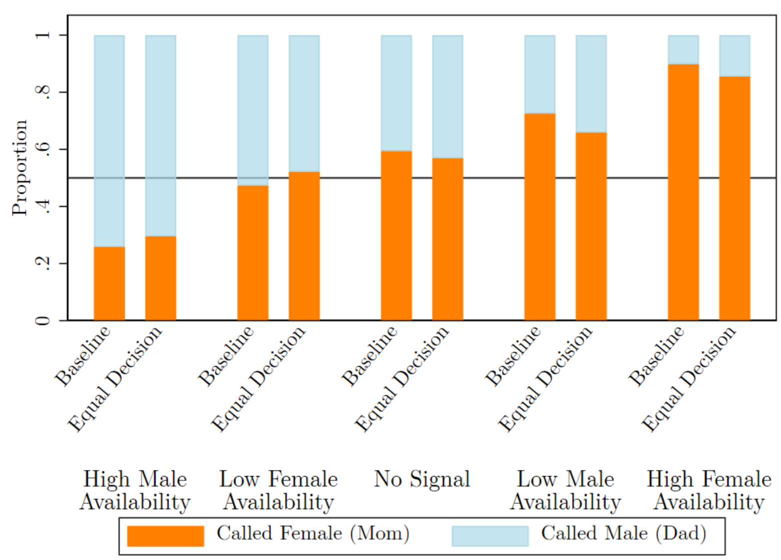
	(1) High Male	(2) Low Female	(3) No Signal	(4) Low Male	(5) High Female
			Panel A.i: Baseline All Outcomes		
Called Female	0.05	0.10	0.12	0.15	0.19
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Called Male	0.16	0.11	0.08	0.06	0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No Call	0.79	0.79	0.79	0.79	0.78
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	7075	5931	5612	5700	6153
			Panel A.ii: Baseline Conditional on Calling		
Called Female | Call	0.26	0.47	0.59	0.73	0.90
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Called Male | Call	0.74	0.53	0.41	0.27	0.10
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	1483	1216	1158	1190	1335
			Panel B.i: Equal Decision All Outcomes		
Called Female	0.06	0.10	0.11	0.12	0.20
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Called Male	0.14	0.09	0.08	0.06	0.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No Call	0.80	0.81	0.81	0.81	0.77
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
Observations	5170	5558	6569	6755	6268
			Panel B.ii: Equal Decision Conditional on Calling		
Called Female | Call	0.30	0.52	0.57	0.66	0.85
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Called Male | Call	0.70	0.48	0.43	0.34	0.15
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Observations	1052	1071	1219	1271	1433

Notes: Standard errors are in parentheses. Observations are weighted so that there are 50% of emails from a female parent and 50% from a male parent and so that all message types have equal weighting.

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 Baseline and Equal Decision variations. Panel (a) represents three outcomes from 60,791 decision-makers, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 11,713$). Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent) within each Variation-Treatment cell (e.g., each bar). See Table 1 for sample size by message and standard errors. See Figures B.2 and B.3 for the total number of no calls, calls to female parents, or calls to male parents by message.

imbalances from existing relationships. Second, the type of inquiry in our messages is not a stereotypical male or female question. Our survey evidence suggests that external decision-makers would exhibit an even stronger bias toward calling female parents if they needed to call a parent to pick up a sick child, discuss allergies, or help with a bake sale.

We explore differences by domain in Section 5.3. However, joining an extracurricular team or paying additional fees (especially at a public school) is 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 significant interruption to a person's day, in contrast to less time-intensive and more flexible requests about an extracurricular team or school fees. As such, we believe that the inequality we document—where the domain is neutral, there are no pre-existing relationships, there is no verbal signal about which parent to contact, and there is no imbalance in the non-verbal signals inherent in who sends the email—is a lower bound on the inequality in external demands from schools on mothers versus fathers.

It is important to note that parents face external demands from many sources, not only from schools. We survey workers in a wide variety of jobs who interact with children and their parents and find that there is a mother preferences in each of the nine domains with which workers self-identify (see panel (b) of Figure 1 and Appendix L.2 for details). Thus, the inequality in external demands from schools that we document compounds across many domains, further exacerbating the impact on job market outcomes for mothers.

While our primary analysis focuses on the first call, we find similar patterns when investigating multiple calls made by the same principals (Figures B.2 and B.3). Conditional on calling, just over half of the principals in our sample make more than one call, with an average principal making 1.7 calls. Principals who make only one call are far more likely to call the mother than the father (about two-thirds to mothers versus one-third to fathers). For those who make a call, only about 40% of those who call the mother first then try the father, while over 50% of those who call the father first then try the mother. The rate of two calls to the mother in a row is double the rate of two calls to the father in a row. Overall, this strongly supports our finding that women are disproportionately more likely to field child-related external demands when no information is provided about parents' relative availability.

3.2 Impact of Signals on Gender Inequality

3.2.1 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 the figure that the signals about high and low availability impact the distribution of calls between parents and can either increase or decrease the bias toward calling female parents.

To rigorously assess how the verbal signals affect bias toward calling mothers in comparison to the No Signal message, Figure B.1 visually represents the outcomes from a multinomial logit model like that in Equation 1 (see Table A.1 for more details). We can apply an appropriate transformation to the estimates from this model to decompose the mechanisms for gender inequality into discrimination based on beliefs about availability versus other deterrents, which we discuss in Section 5.

Recall that we randomly vary signals about availability across four messages: High Male, Low Male, High Female, and Low Female. Two of these messages (High Male and Low Female) go against pre-existing gender norms by stating that the father has a lot of availability or the mother has limited availability. Figure 3 shows that these messages cause calls to move away from mothers and toward fathers, which mitigates the gender gap in external demands. The High Male message reverses the inequality so that mothers are now called 26-30% of the time, while the Low Female message moves mothers and fathers close to parity, with mothers getting 47-48% of the calls and fathers the remaining 52-53% (Table 1). In contrast, the remaining two messages, Low Male and High Female, affirm the gender norm that mothers are more available than fathers. We find that they exacerbate the existing inequality by moving more calls toward mothers and away from fathers.

Our results also highlight a striking asymmetry in the effect of informational interventions. Notably, the High Female message results in her being called between 85-90% of the time, which is in contrast to fathers getting 70-74% of the calls under the High Male 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. It is possible that our messages, especially the signals about low availability, could 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 77% to 81% (Table 1 and Figure B.1).

3.2.2 Nonverbal Signals

In our experiment, we randomly vary verbal cues about which parent is more or less available. Our messages have large effects, with the High Female message resulting in about 20% of principals calling the mother versus only about 5% of principals calling the mother in the High Male message. This is a 14 percentage point difference, which reverses the gender inequality in favor of men (Table 1). However, there are also nonverbal 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 cc'd or vice versa. The person sending the email is a nonverbal signal of which parent to contact first.

Pooling across our treatment messages in the Baseline and Equal Decision variations, we find that the no call rate is similar for both male and female senders, suggesting that principals are as likely to respond to an email regardless of the sender's identity (see Table 2 panel AF.i vs. AM.i and panel BF.i vs. BM.i). However, whether the mother or the father sends the email significantly impacts the gender gap in response. Specifically, sending an email from the mother results in the principal calling her 17-18% of the time and calling the father only 3-4% of the time, a 14 percentage point difference. This is similar to the difference we see between our High Female messages, where the mother is called 19% of the time, and High Male messages, where the mother is called 5% of the time. In contrast, sending the email from the father results in the principal calling him 13-14% of the time and calling the mother 6-7% of the time, a 6 percentage point difference (smaller than the difference between our High Female and High Male messages). It is clear that while the sender's identity has a large positive effect on who gets the first call, that effect is not symmetric for mothers and fathers.

Conditional on a call being made, sending the email from the father results in him being called 65-68% of the time (Table 2, Panels AM.ii and BM.ii, Column 1), meaning that external decision-makers are still calling the mother one-third of the time even when she did not send the message. However, when the mother sends the message, 83-86% of the responding principals call her first (Table 2, Panel AF.ii and BF.ii, Column 1), resulting in the father being called less than one-fifth of the time. This highlights a ceiling on fathers' ability to be the primary contact for child-related tasks.

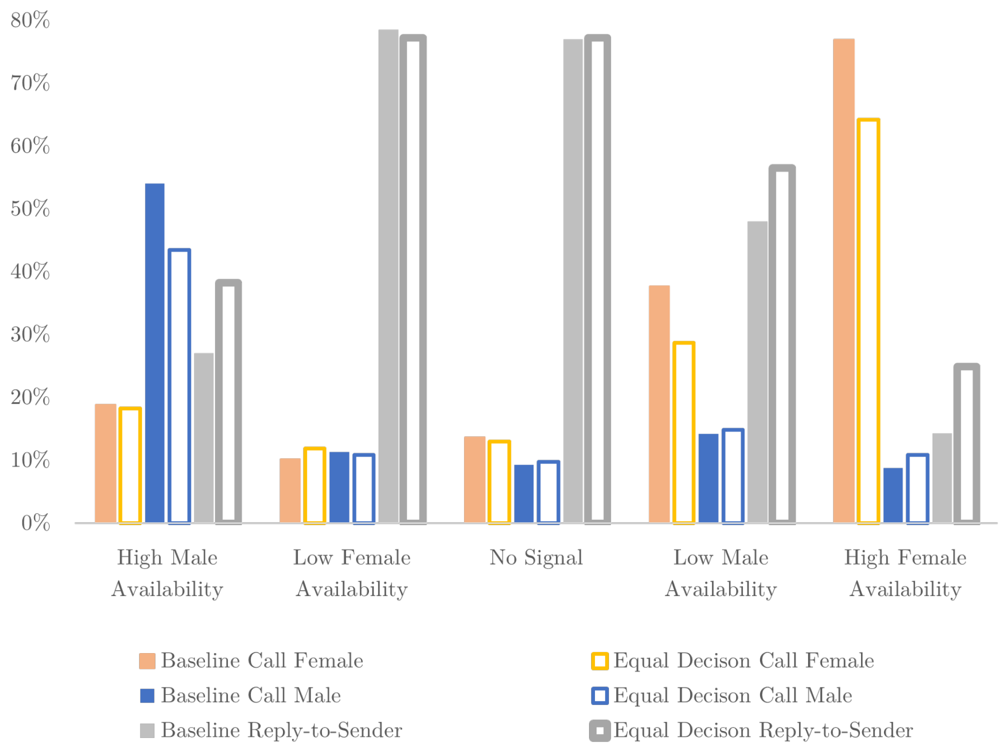
Table 2: **Summary Statistics by Primary Email Sender**

	(1) All Msgs.	(2) High Male	(3) Low Female	(4) No Signal	(5) Low Male	(6) High Female
	Panel AF.i: Baseline Emails Sent by Mother cc'ing Father For All Outcomes					
Called Female	0.18	0.08	0.18	0.20	0.21	0.21
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Called Male	0.04	0.13	0.03	0.01	0.01	0.01
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
No Call	0.79	0.78	0.79	0.79	0.78	0.79
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	15560	3712	2726	3108	2895	3119
	Panel AF.ii: Baseline Emails Sent by Mother cc'ing Father Conditional On Calling					
Called Female | Call	0.83	0.39	0.86	0.98	0.96	0.97
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Called Male | Call	0.17	0.61	0.14	0.02	0.04	0.03
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	3300	801	567	647	626	659
	Panel AM.i: Baseline Emails Sent by Father cc'ing Mother For All Outcomes					
Called Female	0.07	0.02	0.02	0.04	0.10	0.18
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Called Male	0.13	0.18	0.19	0.16	0.11	0.04
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
No Call	0.79	0.80	0.80	0.80	0.80	0.78
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	14911	3363	3205	2504	2805	3034
	Panel AM.ii: Baseline Emails Sent by Father cc'ing Mother Conditional On Calling					
Called Female | Call	0.35	0.12	0.08	0.21	0.48	0.83
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Called Male | Call	0.65	0.88	0.92	0.79	0.52	0.17
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	3082	682	649	511	564	676
	Panel BF.i: Equal Decision Emails Sent by Mother cc'ing Father For All Outcomes					
Called Female	0.17	0.10	0.18	0.18	0.17	0.23
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Called Male	0.03	0.10	0.02	0.01	0.01	0.01
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
No Call	0.80	0.81	0.80	0.81	0.82	0.76
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	15599	2524	3097	3203	3697	3078
	Panel BF.ii: Equal Decision Emails Sent by Mother cc'ing Father Conditional On Calling					
Called Female | Call	0.86	0.50	0.90	0.95	0.96	0.97
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Called Male | Call	0.14	0.50	0.10	0.05	0.04	0.03
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	3084	488	610	600	651	735
	Panel BM.i: Equal Decision Emails Sent by Father cc'ing Mother For All Outcomes					
Called Female	0.06	0.02	0.02	0.03	0.08	0.16
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Called Male	0.14	0.19	0.16	0.15	0.12	0.06
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
No Call	0.80	0.79	0.81	0.82	0.80	0.78
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	14721	2646	2461	3366	3058	3190
	Panel BM.ii: Equal Decision Emails Sent by Father cc'ing Mother Conditional On Calling					
Called Female | Call	0.32	0.11	0.13	0.18	0.40	0.72
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Called Male | Call	0.68	0.89	0.87	0.82	0.60	0.28
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	2962	564	461	619	620	698

Notes: Standard errors are in parentheses. Observations do not have to be weighted in this table by whether the email sender is the mother or father because the panels only show responses to emails from the mother or father. Observations are weighted so that all message types have equal weighting. In columns (1) and (2), the proportions do not always sum to 100% due to rounding, as we have left the output exactly as it came from Stata.

Examining the differences across treatment messages in more detail, we see that three of our messages, No Signal, Low Male, and High Female, result in the mother being called over 95% of the time when she sends the email (Table 2, Panel AF.ii and BF.ii). 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 demands 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.

Figure 4: Calls as a Result of Reply-to-sender vs. Remaining Calls to Female/Male Parent in Baseline and Equal Decision by Treatment



Notes: In this figure, we show the proportion of calls resulting from the decision-maker replying to the primary sender of the email (see footnote 16), alongside the proportion of remaining calls that go to male or female parents conditional on a call being made by treatment in our Baseline and Equal Decision variations ($N = 6,382$ and $N = 6,046$). Note it is difficult to get an equal ratio of Female:Male phone calls in most of the treatments, with the exception of "Low Female Availability." We always cc the other parent in all our emails. Details of "No Call" are shown in Table 2 with very small differences by variation and treatment.

Figure 4 shows the proportion of calls resulting from the decision-maker replying to the primary sender of the email, alongside the proportion of remaining calls that go to male or female parents.[16] It is clear that the reply-to-sender effect is quite large, showing a useful tool for pushing calls from one parent to another. However, something notable about Figure 4 is that almost none of our email treatment pairs result in a 50-50 split in calls to mothers and fathers despite many households reporting they would prefer an equal division of parenting responsibilities. The Low Female treatment is the only one that comes close. This may be because principals are used to the administrative systems employed by most schools and other child-related organizations, which only allow two-parent households to designate a single "Primary Contact." Such a system is likely an artifact of traditional gender norms where one parent focuses on housework while the other focuses on work outside the home. It essentially pushes the household toward a corner solution of always calling mom or always calling dad and is not a viable solution for the ever-increasing share of households desiring a more equitable distribution of child-related tasks.

To investigate the underlying drivers of gender inequality in external demands for parental involvement that we have documented here, we next turn to our theoretical framework. In the next section, we discuss the random utility model that motivates our experimental treatments and allows us to identify structural parameters for decision-makers' beliefs, the importance they place on replying to the parent who reaches out to them, and other deterrents.

4 Theoretical Framework

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

In our specific field experiment, the decision-maker is a school principal, and the task is a discussion about enrolling at the school. However, the model is flexible enough to be applied to different types of decision-makers (e.g., doctors, school teachers, sports coaches, organized religion leaders) and different types of tasks (e.g., picking up a sick child, com-

[16] We randomize which parent is the primary sender of the email (with the other parent CCed). The randomization of who sends the email allows us to quantify the effect of an email being sent from a parent on the likelihood of a reply to that sender. One way to do this is to regress whether a call was made to the female parent on whether the email was sent from the female parent (or analogously regress whether a call was made to the male parent on whether the email was sent from the male parent). Both these regressions give us the height of the gray reply-to-sender bars in Figure 4. The remaining calls are distributed between mothers (orange) and fathers (blue) by multiplying by the proportions in Figure 3 panel b.

municating about health concerns, taking the team on an overnight trip). Furthermore, our model could apply outside of parenting tasks to study many types of demands on a two-person household (e.g., for elder care, home renovations, retirement planning) as long as the central elements are present: one decision-maker, a set of differentiated individuals to contact, and messages that inform key beliefs about the individuals to be contacted.

We lay out a simple economic structure in Section 4.1 to capture the decision-making behavior of school principals when contacting parents. In Section 4.2, we describe the random utility model we have constructed to study this environment. We then explain in Section 4.3 how our experimental variation integrates with the random utility model. Section 4.4 shows how we use the model to identify and estimate its structural parameters, most notably the parameters for principals' beliefs and the other deterrents they face to calling parents. Section 4.5 outlines key testable hypotheses of interest. Appendix G contains additional model details as well as all proofs. It is useful to note here that Appendix G.1 summarizes all model-related notation.

4.1 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 $j \in \{m, f, n\}$. We assign a decision of n to decision-makers who do not make a call by our exogenously-determined experiment end date. The observables in our experiment are then (1) the choice $y_i \in \{m, f, n\}$ for each decision-maker, (2) the characteristics of the alternative that is shown to each decision-maker, and (3) which parent makes the request.[17]

We assume that decision-makers potentially face different costs, c_i, of making a phone call and this cost does not depend on which parent is called. For instance, some may have inferior technology or be busier than others. We also assume that decision-makers potentially perceive different benefits and costs from choosing different alternatives, and that these are made up of three components: the decision-maker's belief about the value of a response from each parent, the decision-maker's value from calling the parent who initially made contact, and the deterrents they face to calling that alternative.[18] We let $r_{ij}q_{ij}$ denote

[17] In Appendix G.5, we extend the model to incorporate the characteristics of the decision-makers.

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

decision-maker i's subjective valuation of a response from alternative j, where r_{ij} is the belief about responsiveness and q_{ij} is the belief about j's desire for equal decision-making within the household. We let s_{ij} be the value the decision-maker derives from calling the person who reached out to them. Finally, we denote by δ_{ij} any other deterrents to calling alternative j. We assume that each decision-maker i knows c_i, s_{ij} and δ_{ij}, has beliefs over r_{ij} and q_{ij}, and is risk neutral.[19]

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

$$U_{ij} = \mathbb{E}(r_{ij}q_{ij}) + s_{ij} - \delta_{ij} - c_i, \tag{2}$$

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

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

Under this normalization and in our context of choice between calling either of two parents or calling neither parent, decision-maker i calls neither parent if both $U_{im} < 0$ and $U_{if} < 0$; calls the female parent if $U_{if} \geq 0$ and $U_{im} \leq U_{if}$; and calls the male parent if $U_{im} \geq 0$ and $U_{if} < U_{im}$.[20]

We can think of a decision-maker's choice between the three alternatives as having two parts: whether to make a call and which parent to call if they are going to make a call. The cost, c_i, does not affect the decision of which parent to call because the decision-maker incurs the same cost regardless of which parent they call. The cost plays a central role in deciding whether to make a call, whereas the choice of which parent to call depends only on the differences in beliefs, the value of replying to the person who sends the email, and

[19]In Appendix G.7, we discuss relaxing the assumption of risk neutrality. Note, in a previous version of the paper, we presented a slightly different version of the model, which we discuss in footnote 13.

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

other deterrents. In order to cleanly identify the parameters of interest, we need to take into account both the decision of whether to make a call and which parent to call, so we need to include the c_i parameter even if it is not of direct interest.

4.3 Experimental Manipulation of Beliefs

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

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

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

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

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

We next describe how decision-maker i updates their beliefs after receiving a signal about parental availability. To keep the notation simple, we focus without loss of generality on how the belief about the female parent is updated, and the case where the prior belief $\bar{q}_j$ equals one.[21] We then have decision-maker i's posterior means for the responsiveness of

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

parent j as

$$\widetilde{rq}_{if}^{F} = \lambda_f^F \bar{r}_f + (1 - \lambda_f^F) x_{if}, \quad \lambda_f^F = \frac{1/\omega_f^2}{1/\omega_f^2 + 1/\sigma_f^2} \tag{3}$$

$$\widetilde{rq}_{if}^{M} = \lambda_f^M \bar{r}_f + (1 - \lambda_f^M) \rho_f x_{im}, \quad \lambda_f^M = \frac{1/\omega_m^2}{1/\omega_m^2 + 1/\sigma_f^2}. \tag{4}$$

$\widetilde{rq}_{if}^{F}$ is the updated belief about the female parent after a signal about the female parent, while $\widetilde{rq}_{if}^{M}$ is the updated belief about the female parent after a signal about the male parent. That is, there are two reasons that decision-maker i's belief about the female parent would be updated: directly via a signal about the female parent or indirectly via a signal about the male parent.[22]

Substituting the updated beliefs into Equation 2 gives us the full model equations, which we rearrange into a reduced form that can be estimated directly from our data on which parent sends the email, which availability message is sent, and who the principal calls.

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

4.4 Identifying the Structural Parameters

If we only send signals about parents' availability as discussed above,[23] we will be able to cleanly identify $\bar{r}_f$ and $\bar{r}_m$ as well as the reply-to-sender parameters for each treatment, s_j^t for $t \in \{\boldsymbol{noSignal}, \boldsymbol{highFemale}, \boldsymbol{lowFemale}, \boldsymbol{highMale}, \boldsymbol{lowMale}\}$. We will not, however, be able to identify the other deterrent parameters or the updating parameters. The problem is that the effects of beliefs about parents' desire for equal decision-making will be absorbed into these parameters.

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

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

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

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

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

This would be a standard random utility model if our reduced-form parameters did not vary across the j choices. However, having intercepts and slopes that vary across alternatives is crucial to learning about how experimental manipulation impacts the choices of decision-makers. Fortunately, the model's structure allows us to identify these intercepts and slopes. Appendix G.2 provides intuition for, and proof of, the identification of the reduced-form parameters. We achieve identification by (1) using the proportions of signal-outcome-sender triplets in the data where there are two distinct signals about each alternative $j \in \{f, m\}$ and (2) imposing known cardinal values for each signal. Specifically, we send both positive and negative signals about each parent's availability and assume the values are 1 and -1.[24] We assume the value of the signal about the desire for equal decision-making is 1 to match the value for the high availability treatment. Appendix G.3 shows that, with these assumptions, the identification of the structural parameters follows directly from the identification of the reduced-form parameters.

[24] Appendix G.6 discusses the robustness of the results to changes in the assigned values of the signals and/or their symmetry.

4.5 Testable Hypotheses

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

The structural parameters identified in Section 4.4 allow us to learn about the sources of this inequality. It may be that decision-makers believe that the expected value of a response from a female parent is higher than the expected value of a response from male parent; we find support for this mechanism if $\bar{r}_f \bar{q}_f > \bar{r}_m \bar{q}_m$. It is also possible that decision-makers face larger deterrents to calling male parents than to calling female parents; we find support for this hypothesis if $\bar{\delta}_m - \bar{\delta}_f > 0$. We examine these questions in the following section.

5 Drivers of the Gender Inequality

Our theoretical model, described in Section 4, allows us to investigate the drivers of the gender inequality we observe in the Baseline No Signal message. Candidate drivers are the decision-maker's beliefs about the value of a response from parents, following the norm of calling the person who sends the message, or other deterrents. In the US, mothers are more likely to be stay-at-home parents than fathers (US Census Bureau, 2022). This general statistical information could lead decision-makers to believe that responses from mothers will provide higher expected value and, as such, will bias decision-makers toward making more external demands of women. In Appendix L.1, we show that these types of decision-makers indeed report that they prefer to contact mothers because they believe mothers are more responsive, but also because they believe mothers are likely to be the primary contact about child-related topics.

Beyond responsiveness, there may be other deterrents affecting decision-makers' choice to call a parent of a certain type. For example, they may prefer talking to mothers because mothers 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 begin by addressing the importance of controlling for the reply-to-sender motive. Because roughly half of our emails are sent from the female parent and half are sent from the male parent, this effect *cannot* drive the gender inequalities in our data. That is, who sends the email impacts this inequality in observational data, but we have experimentally controlled for that by creating balance in which parent sends the email. Another way to think about this is if there were a natural way to send the email from a neutral third party, our results below would not change.

However, when investigating the underlying mechanisms, we may want to consider the potential impact of who sends the emails. The effect of the reply-to-sender motive varies across treatments;[25] we focus here on the effect in the No Signal treatment both because it is an upper bound and because it is most straightforward to think about the effect when it is not interacted with signals about the value of a response. In the No Signal treatment, the utility gain from calling the parent who sends the email is 2.51. For the case when the female parent sends the email, there is a gain of 0.791 from calling the mother and a penalty of 1.722 from calling the father relative to calling neither parent. The result is symmetric when the male parent sends the email. The utility difference is both economically and statistically significantly different from zero ($Prob > chi2 = 0.000$ derived from results in Table A.3). We, thus, conclude that which parent sends the email is an important potential driver of who will be contacted. Since this channel is neutralized in our experiment and controlled for in our estimates, we focus on the role of beliefs and other deterrents in driving the inequalities we observe in our data.

We find that our parameter estimate for the expected value of a response from female parents is $\overline{q}_f \overline{r}_f = -0.341$, which is higher than the analogous parameter for male parents $\overline{q}_m \overline{r}_m = -0.968$. This difference is statistically significant ($Prob > chi2 = 0.013$ derived from results in Table A.3), suggesting that principals believe that mothers are more responsive than fathers. We, thus, find strong support for our hypothesis that beliefs about responsiveness are an important driver of gender inequality in external demands for parents' time.

Next, we test if other deterrents can also explain the gender inequality we document. We find that our parameter estimates for the residual term for male parents are greater than that for female parents; that is, $\overline{\delta}_m - \overline{\delta}_f = 0.536$ ($Prob > chi2 = 0.002$).[26] This is direct evidence

[25]The effect of the reply-to-sender motive is the strongest in the No Signal treatment and weakest in the Female High Availability treatment.

[26]Note, in a previous version of the paper, we presented a slightly different version of the model, which we discuss in footnote 13. This older version of the model reported slightly different parameter estimates, but still

that some gender inequality in demand for parental involvement is driven by factors other than beliefs about responsiveness. Since the difference between the belief parameters is roughly equal to the difference between the other deterrent parameters, we can say that the magnitude of the effect of these other deterrents is about the same as the magnitude of the effect of beliefs about parents' responsiveness. Below, we investigate some of the factors that contribute to both the differential beliefs about the value of response from mothers versus fathers as well as to other deterrents.

5.1 Gender Norms

One mechanism that could explain the gender gap in external demands for parental involvement that we document in our experiment is a strong gender norm governing interactions between decision-makers and parents. As prior studies have shown, despite women's considerable gains in education and labor market outcomes in recent years, social norms about gender identity have persisted and still impact a wide range of economic and social outcomes for women, from labor force participation and earnings to marriage formation, fertility, and the division of home production (Bertrand et al., 2015; Kerwin et al., 2022; Jayachandran, 2021; Ashraf et al., 2023). While we do not have a precise measure of the gender norms of the principals or schools in our sample, we use multiple related measures to investigate whether gender norms may be driving some of the gender inequality in our setting.

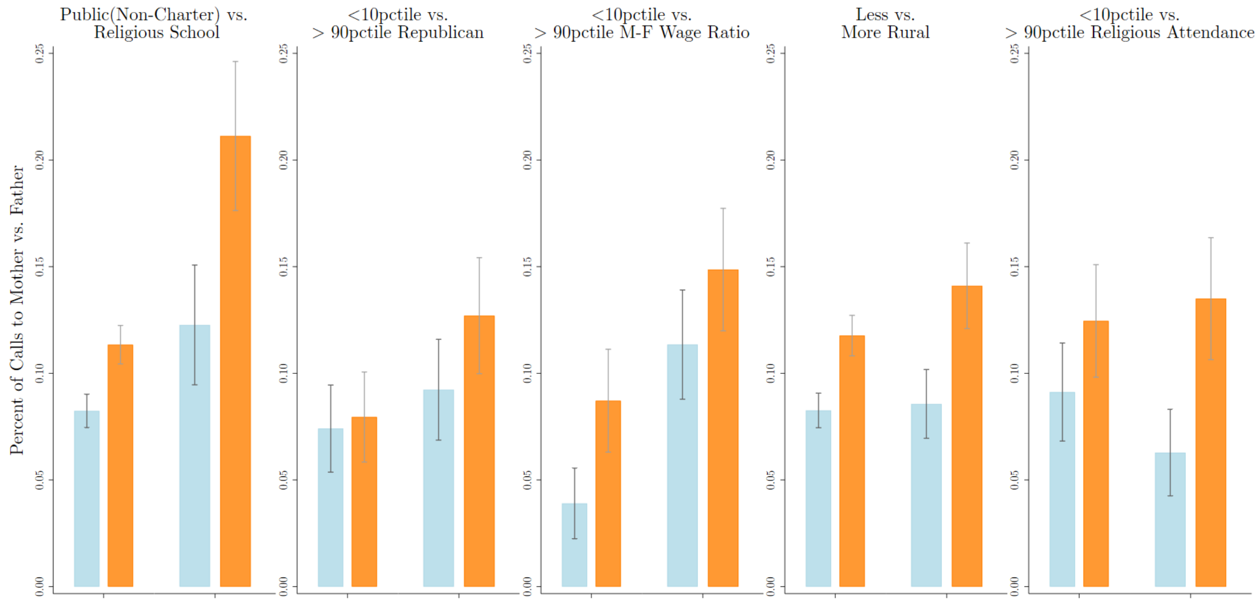
Figure 5 shows that a variety of variables that might be associated with more traditional gender norms are also associated with a higher rate of decision-makers calling the female parent in response to the No Signal message in the Baseline variation. At the most specific level, the school, we observe whether a school is a religious school, which might suggest that it believes in more traditional gender norms. If these gender norms in part drive our results, we would expect greater gender inequality in calls from religious than non-religious schools.[27] This is exactly what we find, especially in the unconditional call proportions. In particular, in the Baseline variation with No Signal, the unconditional call-back rates for religious schools are 21% to mothers and 11% to fathers, versus 12% and 8% for mothers and fathers respectively for non-religious private and public schools (see Table A.4 and a similar pattern in the Equal Decision variation in Table A.5). This difference-in-differences

found that beliefs about responsiveness were greater for female versus male parents, and that $\bar{\delta}_m - \bar{\delta}_f > 0$.

[27] Principals' gender is another dimension where we might see variation in gender norms. However, we find little difference in the patterns by the gender of the principal (Figures D.1 and D.2). While it is theoretically possible that decision-makers forward the email to another person of a different gender, such that we would not capture differences by decision-maker gender, as explained in section K, fewer than 4% of the voicemails left were from someone other than the principal.

is statistically significant ($p = 0.08$).

Figure 5: Differences in Gender Gap by Gender Norm Proxies With No Signal Message in Baseline



Notes: In this figure we show the mean calls to male versus female parents split over proxies for more traditional gender norms (religious school, Republican county, large gender wage gap, more rural, more religious). These are from decision-makers who received our No Signal Message in our Baseline Variation. The details of how these proxies are defined and more details are available in Tables A.4 and A.5. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

We also link our schools to other indicators of gender norms in the county in which the school is located. Specifically, we look at the proportion of Republican voters in the 2016 presidential election, the median wage gap between male and female workers, whether the county is more rural, and whether the county has a higher rate of religious attendance. We find that the proportion of calls to moms is significantly higher in counties with a higher Republican share, larger gender wage gap, and counties that are more rural (see Figure 5 and Tables A.4 and A.5).[28] Note that the number of observations decreases significantly when we compare the gender gap in calls in counties with more traditional versus less traditional gender norms (see Table A.4), resulting in most difference-in-difference estimates being statistically insignificant. However, on net these findings provide strong evidence of

[28] Additionally, we can measure gender norms directly using a sexism index based on data from the General Social Survey (GSS) but these data are only available at the state level. Matching at the state level for an individual school/principal decision makes this measure quite noisy. For example, New York State has a very centrist sexism index, but this masks that New York City is likely relatively non-sexist, while upstate New York may be more sexist. Here, we do not observe the same pattern of greater inequality in calls in more sexist states (Tables A.4 and A.5). We believe this is because measuring norms at the state level is too inexact.

the important role that gender norms play in perpetuating gender inequality in external demands for parents' time.

5.2 Beliefs about Stay-at-Home Mothers

In the US, mothers are significantly more likely to be stay-at-home parents than fathers (US Census Bureau, 2022). 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 in what we call the Full Time variation: "We both work full time." This sentence is meant to shut down the mechanism that the mother is a stay-at-home parent. We sent emails with this message to an additional 9,472 principals (see Appendix E for details by message variations).

Table 3: **Summary Statistics by Variation (All Treatments Combined)**

	Panel A: All Outcomes			
	(1) Baseline	(2) Equal Decision	(3) Full Time	(4) Payments
Called Female	0.124	0.117	0.113	0.100
	(0.002)	(0.002)	(0.003)	(0.003)
Called Male	0.085	0.083	0.077	0.067
	(0.002)	(0.002)	(0.003)	(0.002)
No Call	0.791	0.800	0.810	0.833
	(0.002)	(0.002)	(0.004)	(0.004)
Observations	30471	30320	9472	9808
	Panel B: Conditional on Calling			
Called Female | Call	0.593	0.587	0.594	0.600
	(0.006)	(0.006)	(0.012)	(0.012)
Called Male | Call	0.407	0.413	0.406	0.400
	(0.006)	(0.006)	(0.012)	(0.012)
Observations	6382	6046	1817	1636

Notes: Standard errors are in parentheses. Observations are weighted so that there are 50% of emails from a female parent and 50% from a male parent and so that all message types have equal weighting. Outcomes by the exact message sent within these variations are available in Appendix E.

We would expect fewer calls to mothers in our Full Time variation if beliefs that mothers were more likely to be stay-at-home parents are driving the gender inequality. We do not find evidence of this as shown in Table 3. The rates of calls to mothers and fathers are quite similar in the Full Time variation and the Baseline variation. In the Full Time variation, mothers receive 11.3% of the calls, and fathers receive 7.7% of the calls, which is almost identical to the Baseline variation. Conditional on a call being made, the mother is called 59.4% of the time. In fact, the ratio of calls to mothers versus fathers rises very slightly from 59.3% in the Baseline variation when we include information that shuts down the idea that the mother is a stay-at-home parent.

5.3 Gender Inequality in More Male-Stereotyped Domains

Finally, another factor that could be contributing to the inequality we document is a gender norm about what constitutes a male versus a female domain. In principle, it is possible that both male and female parents are fielding a similar volume of external requests, but certain types of requests are associated with either the female or male domain. Our survey (Appendix L.1) found that, within the school setting, educators stated they most heavily favored calling the mother for a child being sick, for volunteering at a book fair, and when dealing with allergies. While the educators still favored the mother for all other questions, they did so to a lesser degree for requests to volunteer for a career day and to discuss school payments.[29]

To test if fathers are contacted more often in more male-stereotyped domains, we fielded an additional variation of our email messages that stated, "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 parents of either gender, and the differences in call-back rate are driven by emails sent to non-private schools, where perhaps discussion of fees is less common.[30] However, the actual rate of calling mothers versus fathers conditional on a call being made is not statistically significantly different from the Baseline variation at 59.3% (versus 60.0%). Thus, even in the most stereotypical male domain within the school setting, we do not see a shifting of the calls from mothers to fathers.

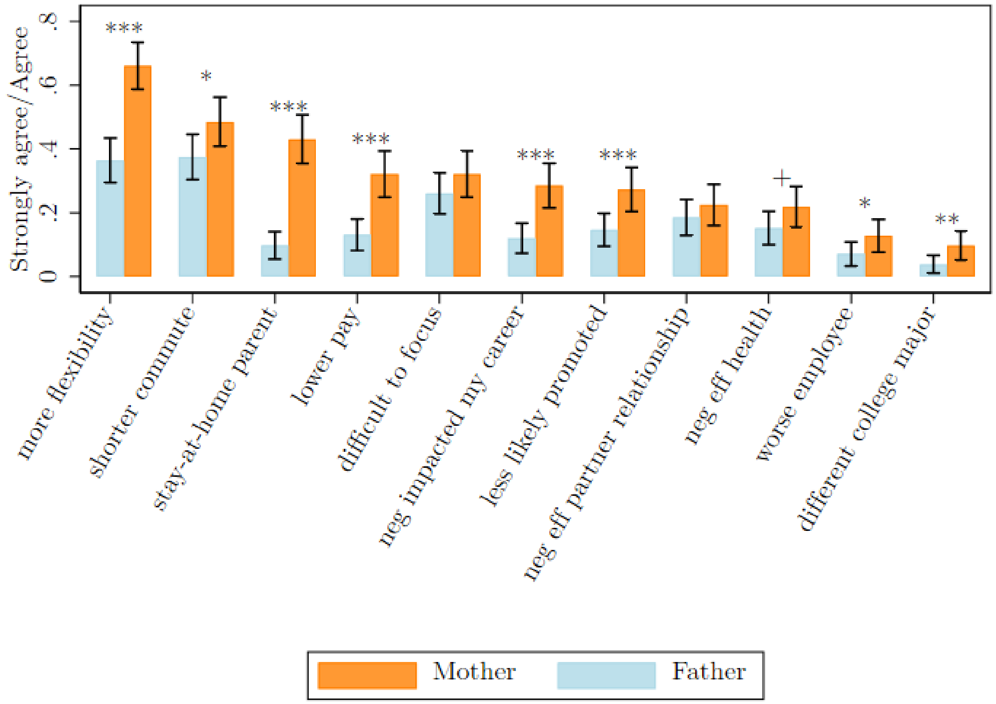
6 Consequences of the Gender Gap in External Demand for Parents' Time

In this section, we investigate the extent to which the gender gap in external demands for parents' time has long-term consequences for the gender gaps in labor market outcomes. We conduct two distinct analyses. First, we administer a survey to mothers and fathers in two-parent households (N = 349, 47% female, see Appendix L.3 for details) asking about specific ways in which child-related external interruptions have impacted their decisions. As reported in Figure 6, we find that across all eleven outcomes, mothers report considerably higher impacts of child-related interruptions on their careers than fathers. For ex-

[29] Prior studies have also found that finances tend to be a more stereotypical male domain (Lin et al., 2022).

[30] In private schools, there is no economically significant change in the No Call rate between our Baseline variation and the one which mentions payments (71% in Baseline vs. 73% in Payments). However, for non-private schools, the No Call rate is 80% in our Baseline variation but 85% when our messages mention payments. All these comparisons are statistically significant.

Figure 6: Changes to Labor Market Choices Due to Child Interruptions



Notes: In this figure, we show the results from a survey of 349 persons who identify as mothers (47%) or fathers (53%) in two parent households with children in the United States. They were asked to indicate how strongly they agreed or disagreed with each of the following statements about whether "child-related interruptions have led me to choose..." or "have led to..." There were five choices: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree. In this figure we show the proportion who stated they either Strong Agree or Agree by gender. We perform one-way t-tests comparing the mean for mothers versus fathers with + $p < 0.10$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

ample, women are more likely than men to say that child-related interruptions have negatively impacted their career trajectory and led them to choose a job that offers lower pay and promotion prospects and allows for more flexibility and shorter commute (all gender differences are economically and statistically significant, $p < 0.01$). These results provide direct evidence that women experience higher career penalties as a result of child-related interruptions and that these interruptions likely contribute to the persistent gender earnings gap.

Notably, we also find that child-related interruptions impact women's educational and labor market participation decisions. Specifically, 43% of women report that child-related interruptions led them to become a stay-at-home parent as compared to only 10% of men ($p < 0.001$). Women are also more likely than men to report choosing their college major in response to anticipating and experiencing child-related interruptions (10% vs. 4%, $p < 0.01$). Finally, we find that that women are also more likely than men to report that child-related interruptions negatively impact their mental and physical health (one-way t-test $p = 0.06$).[31]

[31] We also find women are also more likely than men to report that child-related interruptions negatively

These findings are consistent with prior work documenting that women anticipate labor market effects of motherhood (Kuziemko et al., 2018) and change their employment choices in response to childcare needs (Anstreicher and Venator, 2024).

The second analysis we conduct to quantify the labor market consequences of child-related external demands follows the methodology of Cubas et al. (2021). Using ATUS data[32] and restricting responses to full time working adults with children in two-parent households, we replicate their finding that 35% of women experience a household interruption on a typical workday versus 20% of men (Table M.1). Next, we extend the Cubas et al. (2021) calculations to explore the intensive margin of how fielding a larger proportion of household interruptions might negatively impact wages. The ATUS data allows us to observe the average number of hours per workday parents spend on these interruptions, which are 0.12 for fathers and 0.17 for mothers (Table M.1). That is, in total there are about 0.29 hours of interruptions in a workday for full-time working parents who live with a spouse, and those are split with mothers fielding 58% of those hours, and fathers the remaining 42%. Notably, this is almost exactly the same split as what we observe in our Baseline No Signal experimental treatment. We then compute the wage penalty that is associated with each additional hour of interruption, and find that it corresponds to a 3.4% wage decline (Table M.2). Thus, while each interruption in isolation may not be very time consuming, in combination, these disruptions are associated with economically significant reductions in wages.

Furthermore, we observe that with the current allocation of interruption hours, men's weekly wages are about 7.7% higher than women's (Table M.1). This suggests that our "No Signal" Baseline or Equal Decision variations (where mothers received 59% and 57% of calls respectively) are associated with a 7.7% gender wage gap in households with two full time working parents living together. Taken together, these findings underscore important labor market costs that women incur as a result of bearing the brunt of child-related interruptions from external decision-makers like schools. The inequality in external demands for parental involvement has real consequences to women's labor market outcomes and reinforces existing gender inequalities in society.

---

impact women's ability to focus (one-way t-test $p = 0.11$) as well as their partner relationships (one-way t-test $p = 0.18$), although these differences do not reach statistical significance at traditional levels.

[32] We restrict the data to 2003-2018 to avoid COVID-related issues with coding workdays and to closely match the work of Cubas et al. (2021).

7 Conclusion

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

We document a prominent gender gap in responses. Conditional on receiving a call, 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 the availability of fathers mitigates this inequality and causes mothers to be called less than half the time. However, there is a striking asymmetry in the effects of our informational interventions. Specifically, even when fathers explicitly signal their availability, mothers are still called 26% of the time. In contrast, signals that reinforce stereotypes about mothers being more available cause them to receive 90% of the calls. Notably, even when the email comes from the father *and* he signals his availability, 12% of the calls are still directed to mothers. In contrast, fathers receive only 3% of the calls when mothers send the email and signal that they are available. This underscores a ceiling on the degree to which informational signals can mitigate gender inequality in external demands for parental involvement.

The gender inequality in external demands for parents' time persists even when we account for the non-verbal signal of parents' availability, the identity of the email sender. While sending the email from the father significantly raises the share of calls to fathers, such a system does not offer a solution for the increasing number of households trying to attain a 50-50 allocation, as it effectively pushes the calls to one parent. Additionally, in our survey of parents, we find that even when parents designate the father as the primary contact, organizations will still call mothers about 50% of the time, as compared to only 27% for non-primary contact fathers.

Our theoretical model allows us to disentangle the mechanisms underlying any differential demand for parental involvement 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 each parent's availability and/or desire for equality, while the other factors are measured as a residual term in our model. We find that both beliefs about mothers being more responsive than fathers and differences in the residuals drive the gender inequality in our setting. We test several potential deterrents, including beliefs about mothers being more likely to be stay-at-home parents and the role of gender norms and find evidence that gender norms are, in part, responsible for the gender gap in external demands for parental involvement.

The gender gap in external demands on parents' time can have detrimental and persistent effects on women's career trajectories. Consistent with prior studies, we link more frequent workday interruptions for women versus men to inequalities in a wide range of important economic outcomes, including earnings, occupational choice, human capital accumulation, and promotions. Furthermore, if women are disproportionately shouldering child-related, caregiving, and household tasks, they incur substantial personal costs, including physical and mental health. Investigating the source of these inequalities and documenting that external demands in part drive them 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. To mitigate this gap, it is essential for parents to signal the availability of fathers and their desire for equality *and* for schools to foster more equitable parental involvement.

Notably, the patterns we document likely 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 women are disproportionately more likely to experience interruptions on a daily basis.[33] The gender inequality in physical housework, for example, has remained largely unchanged since the mid-1990s, with men spending about half as much time on housework as women in similar households (Bianchi et al., 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.[34] While these inequalities are more difficult to measure directly, our findings shed light on potential policies to mitigate these gender gaps.

[33] In our survey, we find that women are significantly more likely to be contacted by external decision-makers across a wide range of child-related domains, from doctors' offices to extracurricular sports coaches to religious leaders (see panel (b) of Figure 1). Other studies have documented gender inequality in taking on care-taking in larger samples (Wikle and Cullen, 2023; Bianchi et al., 2006; Boye, 2015; Daly and Groes, 2017; Daminger, 2019; Bertrand et al., 2015; Charmes, 2019).

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

Since the interaction that we investigate involves multiple parties, there are many trade-offs to consider in assessing whether the gender gap in external demands is efficient. For example, external decision-makers may have multiple competing objectives, including getting the most useful response and involving the most diverse set of parents. Disproportionately calling mothers may, thus, be inefficient depending on the specific objectives being maximized. From the perspective of the parents, survey evidence suggests that they prefer a more equal distribution of child-related external demands, and the existing skew towards mothers contributes to both intra-household and labor market inefficiencies. Even if we assume that men and women *on average* have different comparative advantages, there is a distribution of skills within each gender. This implies that households differ from the population average, resulting in dead-weight loss of one-size-fits-all policies due to household inefficiencies. Reducing the restrictions placed on households by institutions would, therefore, lead to a more optimal outcome. Moreover, the skew towards mothers may be welfare-harming for children, given the evidence that children benefit from having both fathers and mothers involved (Pleck, 2007; Nakata, 2023). In Appendix I, we discusses efficiency considerations in more detail.

Finally, while we have documented that mothers are significantly more likely to field external demands than fathers, we do not observe who actually completes the task after being contacted. In principle, it is possible for mothers to outsource the task to their partners. In our survey of parents, for example, we find that parents report doing so quite often, albeit mothers significantly less than fathers (40% vs. 61% respectively when asked about organizations their children attend). Mothers are also twice as likely as fathers to say that outsourcing the task to their partner is disruptive to their day and that they still have to be involved in the task even after asking their partner for help (60% for women vs. 43% for men). This demonstrates that parents exert effort and incur the associated communication and disruption costs in order to achieve the balance of child-related labor they seek for their households, and these costs are higher for women. An approach that creates the balance that parents desire in the first instance would be far superior, as it would avoid significant extra costs to households and help institutions external to the households resolve issues faster.

References

Adams-Prassl, Abi, Kotaro Hara, Kristy Milland, and Chris Callison-Burch, "The gender wage gap on an online labor market: the cost of interruptions," *Review of Economics and Statistics*, 2023.

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

Agan, Amanda Y, Bo Cowgill, and Laura K Gee, "The Tradeoffs of transparency: Measuring inequality when subjects know they are in an experiment," *Working Paper*, 2023.

Aguiar, Mark and Erik Hurst, "Measuring Trends in Leisure: The Allocation of Time Over Five Decades," *The Quarterly Journal of Economics*, 2007.

Ahmed, Ali, Mats Hammarstedt, and Karl Karlsson, "Do Swedish schools discriminate against children with disabilities?," *Working Paper*, 2020.

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

Albanese, Andrea, Adrián Nieto, and Konstantinos Tatsiramos, "Job Location Decisions and the Effect of Children on the Employment Gender Gap," *CESifo Paper 9792*, 2022.

American Family Survey, "American Family Survey," "https://csed.byu.edu/american-family-survey" 2022.

Amuedo-Dorantes, Catalina, Miriam Marcén, Marina Morales, and Almudena Sevilla, "COVID-19 School Closures and Parental Labor Supply in the United States," IZA Discussion Papers 13827, Institute of Labor Economics (IZA) October 2020.

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

Angelov, Nikolay, Per Johansson, and Erica Lindahl, "Parenthood and the gender gap in pay," *Journal of labor economics*, 2016, *34* (3), 545–579.

Anstreicher, Garrett and Joanna Venator, "To Grandmother's House We Go: Informal Childcare and Female Labor Mobility," *Working paper*, 2024.

APA, "American psychological association," 2021.

Arrow, Kenneth, "The theory of discrimination,," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton: Princeton University Press, 1973, pp. 3–3.

Ashraf, Nava, Oriana Bandiera, Virginia Minni, and Victor Quintas y Martinez, "Gender Roles and the Misallocation of Labour Across Countries," *Working Paper*, 2023.

Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart, "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability," *American Economic Review*, March 2017, *107* (3), 714–47.

Becker, Gary S, *The economics of discrimination*, University of Chicago press, 1957.

Belkin, Lisa, "Dads in the PTA," *New York Times Mortherlode Blog*, 2009.

Bergman, Peter and Isaac McFarlin Jr, "Education for all? A nationwide audit study of school choice," *Working Paper*, 2018.

Bertrand, Marianne and Esther Duflo, "Field experiments on discrimination," *Handbook of economic field experiments*, 2017, *1*, 309–393.

_ **and Sendhil Mullainathan**, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American economic review*, 2004, *94* (4), 991–1013.

_ **, Emir Kamenica, and Jessica Pan**, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

Bianchi, Suzanne M, John P Robinson, and Melissa A Milkie, *Changing rhythms of American family life*, Russell Sage Foundation, 2006.

_ **, Liana C Sayer, Melissa A Milkie, and John P Robinson**, "Housework: Who did, does or will do it, and how much does it matter?," *Social forces*, 2012, *91* (1), 55–63.

BLS, "Bureau of Labor Statistics," 2021.

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

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate Statistical Discrimination: An Identification Problem," Review of Economics and Statistics forthcoming.

Bohren, J Aislinn, Peter Hull, and Alex Imas, "Systemic discrimination: Theory and measurement," Technical Report, National Bureau of Economic Research 2022.

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

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais, "'Acting wife': Marriage market incentives and labor market investments," *American Economic Review*, 2017, *107* (11), 3288–3319.

Cantet, Natalia, Brian Feld, and Monica Hernandez, "Is there discrimination against children of same-sex households?," *Working Paper*, 2022.

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

Census, "Census Bureau Survey Explores Sexual Orientation and Gender Identity," "https://www.census.gov/library/stories/2021/11/census-bureau-survey-explores-sexual-orientation-and-gender-identity.html" 2021.

—, "Most Kids With Parent in Same-Sex Relationship Live With Female Couple," "https://www.census.gov/library/stories/2022/07/most-kids-with-parent-in-same-sex-relationship-live-with-female-couple.html" 2022.

Charmes, Jacques, "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*, 2019.

Charness, Gary, Anya Samek, and Jeroen van de Ven, "What is considered deception in experimental economics?," *Experimental Economics*, 2022, *25* (2), 385–412.

Ciasullo, Ludovica and Martina Uccioli, "What Works for Working Mothers? A Regular Schedule Lowers the Child Penalty," *Working Paper*, 2023.

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

Conger, Rand D, Katherine J Conger, and Monica J Martin, "Socioeconomic status, family processes, and individual development," *Journal of marriage and family*, 2010, *72* (3), 685–704.

Cortes, Patricia and Jessica Pan, "Children and the Remaining Gender Gaps in the Labor Market," *Working Paper*, 2021.

Couch, Kenneth A, Robert W Fairlie, and Huanan Xu, "The evolving impacts of the COVID-19 pandemic on gender inequality in the US labor market: The COVID motherhood penalty," *Economic Inquiry*, 2022, *60* (2), 485–507.

Council, National Research, "Measuring Racial Discrimination," 2004.

Cowan, Benjamin W, Todd R Jones, and Jeffrey M Swigert, "Parental and Student Time Use Around the Academic Year," Technical Report, National Bureau of Economic Research 2023.

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

Cubas, G., C. Juhn, and P. Silos, "Coordinated Work Schedules and the Gender Wage Gap," *Economic Journal*, 2022.

Cubas, German, Chinhui Juhn, and Pedro Silos, "Work-Care Balance over the Day and the Gender Wage Gap," in "AEA Papers and Proceedings," Vol. 111 2021, pp. 149–53.

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

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

Diaz-Serrano, Luis and Enric Meix-Llop, "Do schools discriminate against homosexual parents? Evidence from a randomized correspondence experiment," *Economics of Education Review*, 2016, *53*, 133–142.

Duchini, Emma and Clémentine Van Effenterre, "School schedule and the gender pay gap," *Journal of Human Resources*, 2022.

Erosa, Andrés, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson, "Hours, Occupations, and Gender Differences in Labor Market Outcomes," *American Economic Journal: Macroeconomics*, 2022, *14* (3), 543–90.

Garcia, Kairon Shayne D and Benjamin W Cowan, "The impact of school and childcare closures on labor market outcomes during the COVID-19 pandemic," Technical Report, National Bureau of Economic Research 2022.

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

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

Hansen, Benjamin, Joseph Sabia, and Jessamyn Schaller, "Schools, Job Flexibility, and Married Women's Labor Supply," *Working Paper*, 2022.

Heggeness, Misty L, "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*, 2020, *18* (4), 1053–1078.

Hermes, Henning, Philipp Lergetporer, Philipp Peter, Fabian Mierisch, and Simon Wiederhold, "Males Should Mail? Gender Discrimination in Access to Child Care," in "AEA Papers and Proceedings" 2023.

Islam, Asad, Debayan Pakrashi, Liang C Wang, and Yves Zenou, "Determining the Extent of Statistical Discrimination: Evidence from a field experiment in India," 2018. Available at SSRN: https://ssrn.com/abstract=3185899.

Jack, Rebecca, Daniel Tannenbaum, and Brenden Timpe, "The Parenthood Gap: Firms and Earnings Inequality After Kids," *Working paper.*

Jayachandran, Seema, "Social Norms as a Barrier to Women's Employment in Developing Countries," *IMF Economic Review*, 2021.

Karpowitz, Chris, Stephen O'Connell, Jessica Preece, and Olga Stoddard, "Strength in Numbers: Gender Composition, Leadership, and Women's Influence in Teams," *Working Paper*, 2023.

Kerwin, Charles, Jonathan Guryan, and Jessica Pan, "The Effects of Sexism on American Women: The Role of Norms vs. Discrimination," *Journal of Human Resources*, 2022.

Kleven, Henrik, "The Geography of Child Penalties and Gender Norms: Evidence from the United States," *Working Paper*, 2023.

_ , Camille Landais, and Jakob Egholt Søgaard, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019, *11* (4), 181–209.

_ , _ , and _ , "Does biology drive child penalties? evidence from biological and adoptive families," *American Economic Review: Insights*, 2021, *3* (2), 183–98.

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

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington, "The mommy effect: Do women anticipate the employment effects of motherhood?," Technical Report, National Bureau of Economic Research 2018.

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

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

Lin, Emily, Joel Slemrod, Evelyn Smith, and Alexander Yuskavage, "Who's on (the 1040) First? Determinants and Consequences of Spouses' Name Order on Joint Returns," *Working Paper*, 2022.

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

Mas, Alexandre and Amanda Pallais, "Valuing alternative work arrangements," *American Economic Review*, 2017, *107* (12), 3722–3759.

_ and _ , "Alternative work arrangements," *Annual Review of Economics*, 2020, *12*, 631–658.

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

Montes, Joshua, Christopher Smith, and Isabel Leigh, "Caregiving for children and parental labor force participation during the pandemic," 2021.

Nakata, Kazuko, "Income, Father's and Mother's Time Investment and Children's Achievement," *Working Paper*, 2023, *0* (0).

NCES, "National Center For Education Statistics," 2021.

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

Oberfield, Zachary W and Matthew B Incantalupo, "Racial Discrimination and Street-Level Managers: Performance, Publicness, and Group Bias," *Public Administration Review*, 2021, *81* (6), 1055–1070.

Offer, Shira, "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," Vol. 29 Wiley Online Library 2014, pp. 916–936.

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

Pager, Devah, "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future," *The ANNALS of the American Academy of Political and Social Science*, 2007, *609* (1), 104–133.

Perez-Vaisvidovsky, Nadav, Ayana Halpern, Reli Mizrahi, and Zhara Atalla, "Fathers Are Very Important, but They Aren't Our Contact Persons: The Primary Contact Person Assumption and the Absence of Fathers in Social Work Interventions," *Families in Society*, 2023, *0* (0), 10443894221145751.

Pertold-Gebicka, Barbara, Filip Pertold, and Nabanita Datta Gupta, "Employment adjustments around childbirth," *IZA Discussion Paper*, 2016.

Pew Research Center, "Raising Kids and Running a Household: How Working Parents Share the Load," 2015.

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

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

Powell, Walter W and Paul J DiMaggio, *The new institutionalism in organizational analysis*, University of Chicago press, 2012.

Price, Brendan M and Melanie Wasserman, "The summer drop in female employment," *Working Paper*, 2022.

Russell, Lauren and Chuxuan Sun, "The Effect of Mandatory Child Care Center Closures on Women's Labor Market Outcomes During the COVID-19 Pandemic," 2020.

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

Scott, W Richard, *Institutions and organizations: Ideas, interests, and identities*, Sage publications, 2013.

Scottland, Fathers' Network, "Why Fathers' Involvement Matters," 2020.

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

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

Speer, Jamin and Fulya Ersoy, "Opening the Black Box of College Major Choice: Evidence from an Information Intervention," *Working Paper*, 2022, *22:666.*

SSA, "1980s Names," "https://www.ssa.gov/OACT/babynames/decades/names1980s.htm" 2022.

___ , "Last Names," "https://namecensus.com/last-names/" 2022.

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

US Census Bureau, "Income and Poverty in the United States: 2019," 2020.

___ , "Number of stay-at-home parents among opposite-sex married couple families with children under 15," 2022.

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

Wikle, Jocelyn and Clara Cullen, "The Developmental Course of Parental Time Investments in Children from Infancy to Late Adolescence," *Social Sciences*, 2023, *12* (2), 92.

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

Appendix: For Online Publication Only

A Appendix Tables

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

	(1) Baseline	(2) Equal Decision
Outcome: Female Call		
High Male	-0.81***	-0.54***
	(0.07)	(0.07)
Low Female	-0.23***	-0.04
	(0.06)	(0.06)
Low Male	0.22***	0.17**
	(0.05)	(0.05)
High Female	0.48***	0.67***
	(0.05)	(0.05)
Outcome: Male Call		
High Male	0.62***	0.60***
	(0.06)	(0.06)
Low Female	0.26***	0.15*
	(0.07)	(0.07)
Low Male	-0.38***	-0.21**
	(0.08)	(0.07)
High Female	-1.31***	-0.83***
	(0.10)	(0.08)
Observations	30,471	30,320

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

Table A.2: **Call vs. No Call by Observable Variables of Schools By Variation**

	(1) All Called	(2) All No Call	(3) Baseline Called	(4) Baseline No Call	(5) Equal Decision Called	(6) Equal Decision No Call	(7) Payments Called	(8) Payments No Call	(9) Fulltime Called	(10) Fulltime No Call
Elementary	0.46	0.50	0.47	0.50	0.46	0.50	0.43	0.51	0.49	0.50
Middle	0.14	0.15	0.14	0.14	0.13	0.14	0.17	0.15	0.14	0.16
High	0.18	0.20	0.18	0.20	0.17	0.19	0.20	0.20	0.17	0.21
Decison-Maker Female	0.57	0.58	0.57	0.58	0.57	0.58	0.56	0.59	0.58	0.58
PublicCharter	0.05	0.06	0.05	0.06	0.05	0.06	0.06	0.06	0.06	0.06
PublicNOTCharter	0.73	0.80	0.74	0.80	0.73	0.79	0.69	0.81	0.73	0.80
Private	0.22	0.14	0.21	0.14	0.22	0.15	0.25	0.14	0.22	0.14
FreeLunch	0.49	0.56	0.49	0.56	0.51	0.55	0.48	0.56	0.50	0.55
White	0.60	0.50	0.60	0.51	0.60	0.50	0.60	0.50	0.60	0.51
Black	0.11	0.16	0.11	0.15	0.11	0.16	0.12	0.16	0.11	0.15
Hispanic	0.19	0.24	0.19	0.24	0.20	0.24	0.19	0.24	0.20	0.24
FemaleEmail	0.50	0.50	0.51	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Observations	15881	64190	6382	24089	6046	24274	1636	8172	1817	7655

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. Decision-Maker Female is whether the decision-maker (the principal) has a first name that is female. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent).

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

	(1) No Call Base	(2) No Call Base
Panel A: Outcome Female Call (vs. No Call)		
Any Signal About Male	-0.22***	
	(0.07)	
x_M (Male Signal Pos/Neg)	-0.43***	
	(0.04)	
Any Signal About Female	0.19**	
	(0.06)	
x_F (Female Signal Pos/Neg)	0.57***	
	(0.04)	
reply-to-sender*HighMale	0.68***	0.68***
	(0.07)	(0.07)
reply-to-sender*LowMale	0.36***	0.36***
	(0.04)	(0.04)
reply-to-sender*HighFemale	0.20***	0.20***
	(0.03)	(0.03)
reply-to-sender*LowFemale	1.01***	1.01***
	(0.07)	(0.07)
reply-to-sender*NoSignal	0.79***	0.79***
	(0.05)	(0.05)
High Male		-0.65***
		(0.09)
High Female		0.76***
		(0.06)
Low Male		0.21**
		(0.07)
Low Female		-0.37***
		(0.08)
Constant	-2.15***	-2.15***
	(0.05)	(0.05)
Panel B: Outcome Male Call (vs. No Call)		
Any Signal About Male	0.76***	
	(0.14)	
x_M (Male Signal Pos/Neg)	0.78***	
	(0.06)	
Any Signal About Female	0.15	
	(0.15)	
x_F (Female Signal Pos/Neg)	-0.52***	
	(0.07)	
reply-to-sender*HighMale	-0.34***	-0.34***
	(0.04)	(0.04)
reply-to-sender*LowMale	-1.46***	-1.46***
	(0.12)	(0.12)
reply-to-sender*HighFemale	-1.13***	-1.13***
	(0.12)	(0.12)
reply-to-sender*LowFemale	-1.04***	-1.04***
	(0.07)	(0.07)
reply-to-sender*NoSignal	-1.72***	-1.72***
	(0.13)	(0.13)
High Male		1.54***
		(0.14)
High Female		-0.38*
		(0.18)
Low Male		-0.03
		(0.18)
Low Female		0.67***
		(0.15)
Constant	-3.31***	-3.31***
	(0.13)	(0.13)
Observations	29363	29363

Notes: Column 1 of this table presents the results of a multinomial logit model using a model like the one in Equation 1. The outcome variable takes three values: no call, call female, or call male. The right-hand side variables are "Any Signal About Male" which takes the value 1 if a message was sent with a signal about the male parent (MaleHigh, MaleLow) and zero otherwise. "Any Signal About Female" takes the value 1 if a message with a signal about the female parent was sent (FemaleHigh, FemaleLow) and zero otherwise. The variable x_M (Male Signal Pos/Neg) takes the value 1 if the MaleHigh message was sent, and -1 if the MaleLow message, 0 otherwise; x_F is defined analogously for messages about female parents. Column 2 of this table presents results that are discussed in Appendix Section G.4. In both modelts there are a series of variables that control for the gender of the sender of the email (male vs. female parent) interacted with the signals about each parent's availability. The variable "reply-to-sender" takes the value 1 if the sender of the email is female, and -1 if the sender of the email is male (recall we always send an email from one parent and CC the other parent). The variables which capture which of our five messages were sent (MaleHigh, MaleLow, FemaleHigh, FemaleLow and NoSignal) are interacted with "reply-to-sender." The right-hand side variables are discussed in Section 4. In this table we present the results with a base case of no call. Observations are weighted so that 50% of emails come from a female parent and 50% from a male parent (always CCing the other parent). + $p <$ 0.10 * $p < 0.05$ ** $p < 0.010$ *** $p < 0.001$

Table A.4: **More vs. Less Traditional Gender Norms Summary Statistics No Signal Message in Baseline Variation**

	(1) Non Religious School	(2) Religious School	(3) Low Repub. County	(4) High Repub. County	(5) Small Wage Gap County	(6) Large Wage Gap County	(7) Less Rural County	(8) More Rural County	(9) Less Religious County	(10) More Religious County	(11) Less Sexist State	(12) More Sexist State
Called Female	0.11	0.21	0.08	0.13	0.09	0.15	0.12	0.14	0.12	0.14	0.13	0.12
Called Male	0.08	0.12	0.07	0.09	0.04	0.12	0.08	0.09	0.09	0.06	0.08	0.07
No Call	0.80	0.67	0.85	0.78	0.87	0.74	0.80	0.77	0.78	0.80	0.79	0.81
Called Female | Call	0.58	0.63	0.52	0.58	0.69	0.56	0.59	0.62	0.58	0.69	0.63	0.61
Called Male | Call	0.42	0.37	0.48	0.42	0.31	0.44	0.41	0.38	0.42	0.31	0.37	0.39
Observations	4755	528	635	580	529	593	4439	1161	606	553	485	607

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

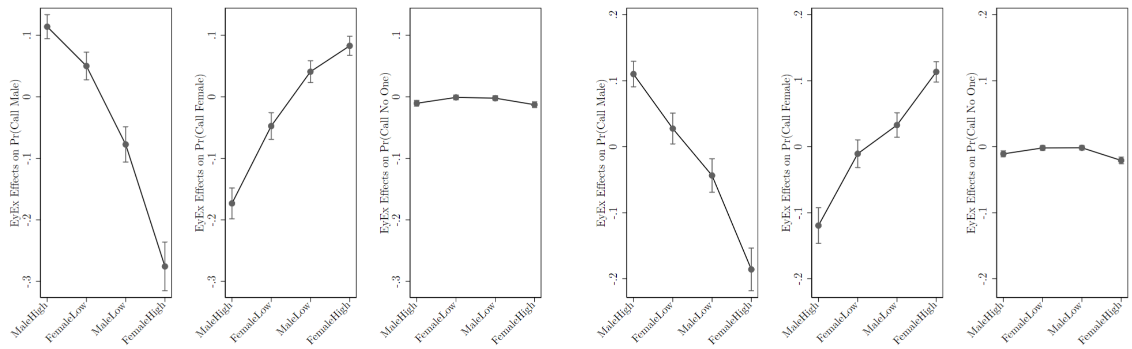
Table A.5: **More vs. Less Traditional Gender Norms Summary Statistics No Signal Message in Equal Variation**

	(1) Non Religious School	(2) Religious School	(3) Low Repub. County	(4) High Repub. County	(5) Small Wage Gap County	(6) Large Wage Gap County	(7) Less Rural County	(8) More Rural County	(9) Less Religious County	(10) More Religious County	(11) Less Sexist State	(12) More Sexist State
Called Female	0.09	0.17	0.08	0.11	0.08	0.12	0.10	0.11	0.11	0.11	0.09	0.09
Called Male	0.08	0.12	0.04	0.09	0.05	0.08	0.08	0.10	0.09	0.09	0.07	0.08
No Call	0.83	0.71	0.88	0.79	0.87	0.80	0.82	0.79	0.80	0.80	0.84	0.83
Called Female | Call	0.55	0.59	0.64	0.55	0.64	0.58	0.58	0.54	0.55	0.54	0.54	0.55
Called Male | Call	0.45	0.41	0.36	0.45	0.36	0.42	0.42	0.46	0.45	0.46	0.46	0.45
Observations	5367	825	853	655	715	630	5209	1345	630	654	607	697

Notes: Variables are defined as in Table A.4.

B Appendix Figures

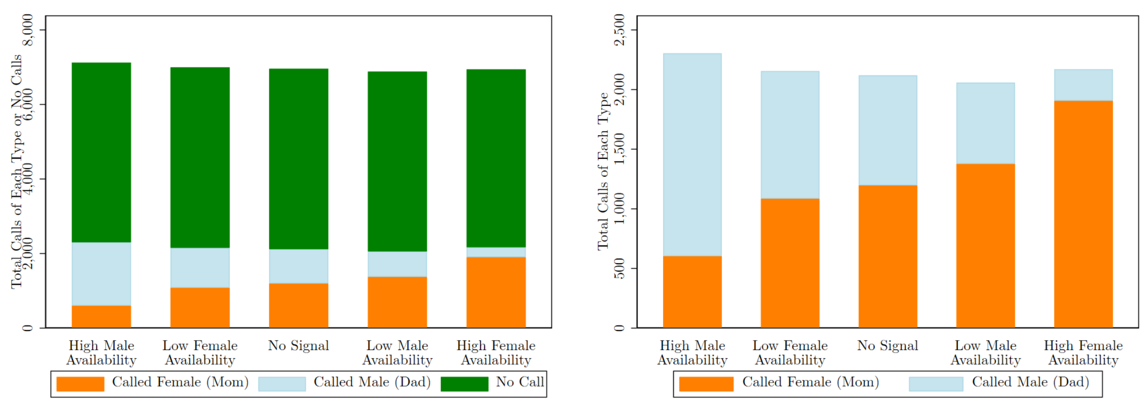
Figure B.1: Effects by Treatment



(a) Baseline Variation (b) Equal Decision Variation

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

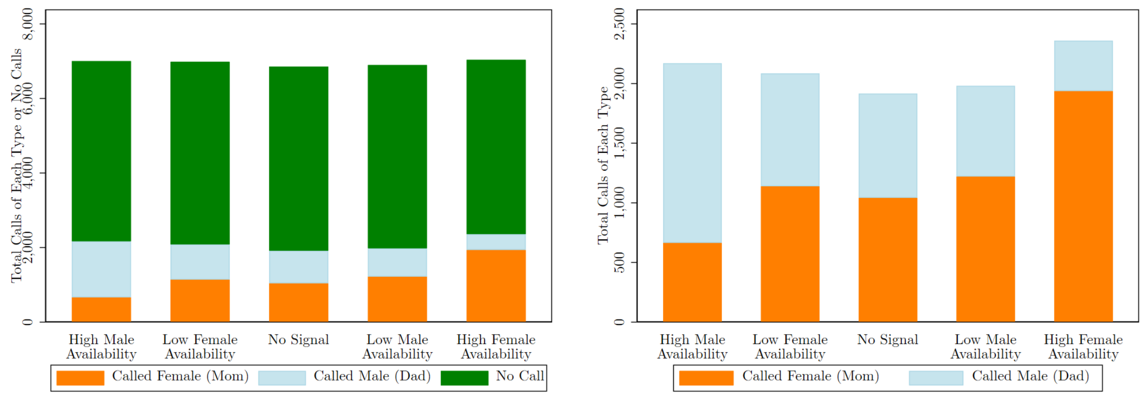
Figure B.2: Outcomes by Treatment in Baseline Variation for Multiple Calls



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

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

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



(a) All Outcomes

(b) Outcomes Conditional on Calling

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

C Balance Tables

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

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

	(1) High Male	(2) Low Female	(3) No Signal	(4) Low Male	(5) High Female
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 preschools). The following variables are known only for non-private schools: FreeLunch, White, Black, Hispanic. DMFemale is whether the decision-maker (the principal) has a first name that is female. Observations are weighted so that 50% of emails are from a female parent and 50% from a male parent.

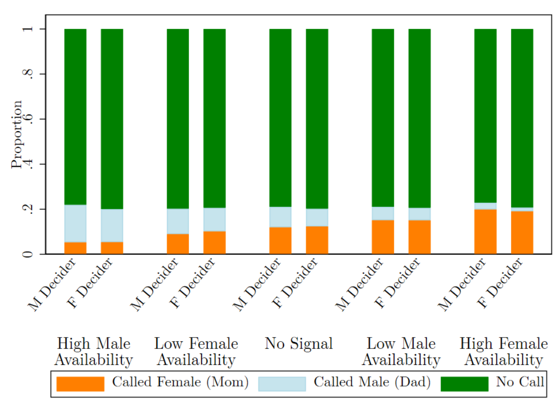
Table C.2: **Balance on Observable Attributes of Schools/Decision-Makers By Treatment In Equal Decision Variation**

	(1) High Male	(2) Low Female	(3) No Signal	(4) Low Male	(5) High Female
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

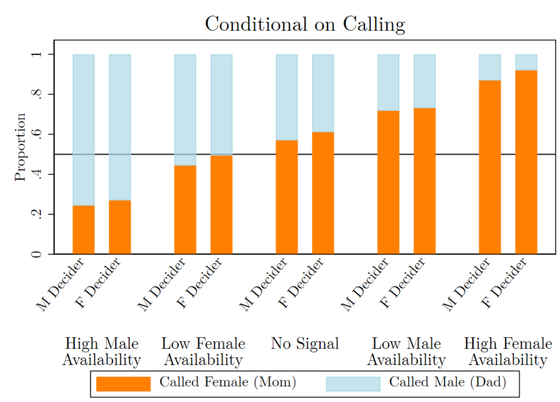
Notes: Notes are the same as those in Table C.1.

D By Decision-Maker Gender

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



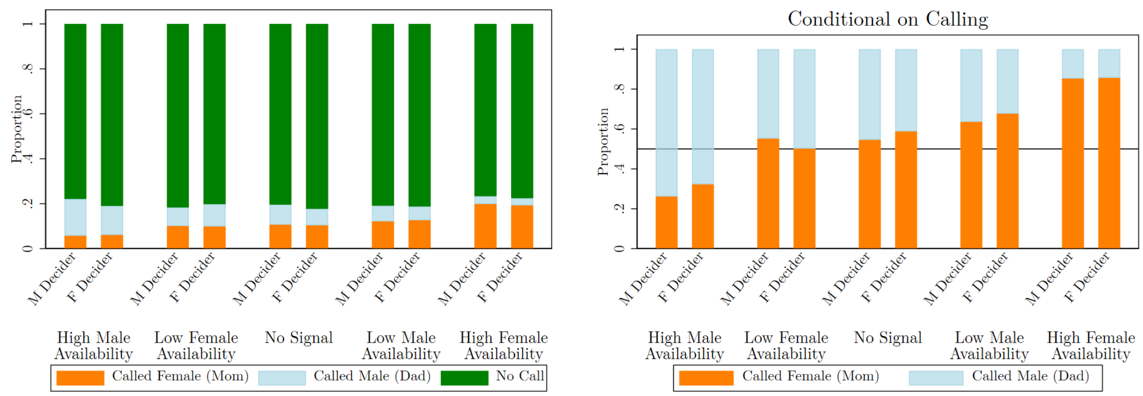
(a) All Outcomes



(b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between Female and Male principals. We predict principal gender based on their name. In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision maker in our Baseline Variation. "M Decider" denotes a male principal and "F Decider" denotes a female principal. Panel (a) represents three outcomes from 30, 471 decision-makers in Main, while panel (b) shows only the choices of those who made a phone call to at least one parent. In Panel B we regress dummy variables for our five messages on a binary variable for whether the female parent was called first or the male parent. If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

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



(a) All Outcomes

(b) Outcomes Conditional On Calling

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

E Variations On Baseline Messages

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

	(1) High Male	(2) Low Female	(3) No Signal	(4) Low Male	(5) High Female
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

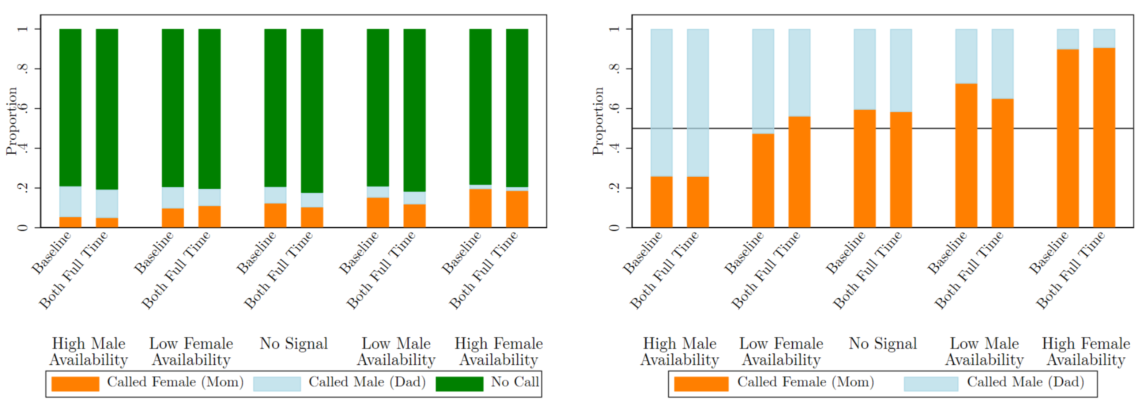
Notes: There is a small proportion of schools which are not Elementary, Middle or High Schools (e.g. K-12 or pre-schools). The following variables are only known for non-private schools: FreeLunch, White, Black, Hispanic. DMFemale is whether the decision maker (the principal) has a first name that is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

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

	(1) High Male	(2) Low Female	(3) No Signal	(4) Low Male	(5) High Female
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: Notes are the same as Table E.1.

Figure E.1: Outcomes By Treatment "Basline" vs. "Full Time" Variations

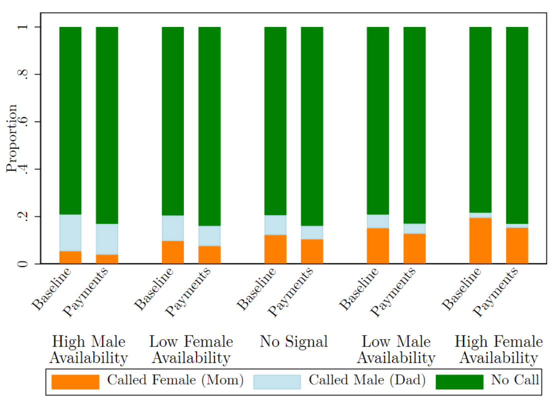


(a) All Outcomes

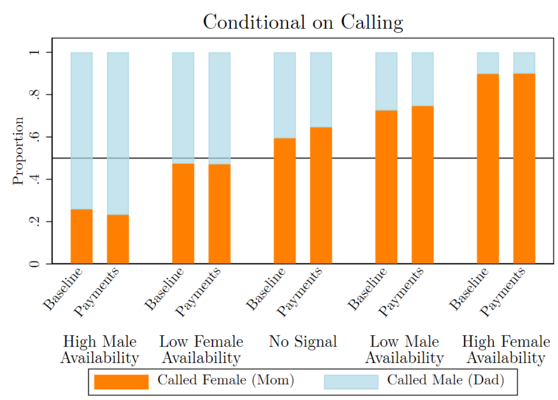
(b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between our "Main" version of our emails and ones that have the addition of a sentence that states "We both work full-time." In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision maker in our Baseline Variation. Panel (a) represents three outcomes from 30,471 decision-makers in Baseline and 9,472 in Full Time, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6382$ in Baseline and 1817 in Full Time). In Panel B we regress dummy variables for our five messages on a binary variable for whether the female parent was called first or the male parent. If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

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



(a) All Outcomes



(b) Outcomes Conditional On Calling

Notes: In this figure we show the differences between our "Main" version of our emails and ones that have the addition of a clauses that states they are "especially interested in discussing school fees and other expenses." In panel (a) we show the proportion of decision-makers choosing to make no call, call the female parent (mom) or the male parent (dad) by the message sent to the decision maker in our Baseline Variation. Panel (a) represents three outcomes from 30, 471 decision-makers in Baseline and 9, 808 in Full Time, while panel (b) shows only the choices of those who made a phone call to at least one parent ($N = 6382$ in Baseline and 1817 in Full Time). In Panel B we regress dummy variables for our five messages on a binary variable for whether the female parent was called first or the male parent. If decision-makers were randomizing which parent they called we would expect the same proportion of calls to male and female parents.

F Example Emails Full Text

Figure F.1: Baseline: No Signal

School Inquiry

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

Dear Principal Gee,

We are searching for schools for our child. Can you call one of us to 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 F.2: Baseline: High Female and Low Female Signal

School Inquiry

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

Dear Principal Gee,

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

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

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

School Inquiry

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

Dear Principal Gee,

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

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

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

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

School Inquiry

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

Dear Principal Gee,

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

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

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

School Inquiry

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

Dear Principal Gee,

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

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

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

G Theory Appendix

G.1 Notation

We provide a summary of our notation as a reference.

Subscripts and superscripts

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

Objects of interest

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

G.2 Identification of Reduced Form Parameters

We first combine the economic structure in Section 4.1 with the random utility model in Section 4.2 and our experimental manipulation in Section 4.3 to derived the reduced form of our model. A summary of the crucial assumptions of those sections follows.

1. Decision maker i chooses from among three alternatives: $j \in \{n, f, m\}$.

2. Decision maker i holds probabilistic beliefs about the probability that alternative j will respond to a phone call, $r_{ij} \sim \mathcal{N}(r_j, \omega_j^2)$.

3. Decision maker i holds probabilistic beliefs about the probability that alternative j will desire to be equally involved in the decision, $q_{ij} \sim \mathcal{N}(\bar{q}_j, v_j^2)$.

4. Each decision maker faces a cost c_i for making a call that is the same for alternatives f and m. c is the population mean of c_i.

5. Each decision maker has a deterrent parameter for calling that varies by alternative.

6. Each decision maker has a preference for responding to the parent who sends the message. This preference may depend on the message. We define the variable s_{ij} to be equal to 1 when the female parent sends the message and -1 when the male parent sends the message and we allow for interactions with each treatment, which we denote of s_{ij}^t.

7. Each decision maker i knows c_i, s_{ij}^t, and δ_{ij}.

8. Decision makers are risk neutral.[1]

9. Expected utility for decision maker i is $\mathbb{E}(U_{ij}) = \mathbb{E}(q_{ij}r_{ij}) + s_{ij} - (\delta_{ij} + c_i)$ for $j \in \{n, f, m\}$ with $\mathbb{E}(U_{in}) = 0$.

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

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

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

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

[1] We have assumed that decision makers are risk neutral with respect to the decision about whether and whom to call. In Appendix G.7, we discuss relaxing this assumption.

We must take a stand on how decision makers will interpret our signals about availability given that their beliefs also contain the desire-for-equal-decision making component. If the signals about availability and the desire for equal involvement do not interact, the beliefs in Expressions 3 and 4 about females become

$$\widetilde{qr}_{if}^{F} = \lambda_f^F \overline{q}_f \overline{r}_f + (1-\lambda_f^F) x_{if}, \quad \lambda_f^F = \frac{1/\omega_f^2}{1/\omega_f^2 + 1/\sigma_f^2} \tag{5}$$

$$\widetilde{qr}_{if}^{M} = \lambda_f^M \overline{q}_f \overline{r}_f + (1-\lambda_f^M) \rho_f x_{im}, \quad \lambda_f^M = \frac{1/\omega_m^2}{1/\omega_m^2 + 1/\sigma_m^2} \tag{6}$$

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

We let w_{ij} be an indicator for sending i a verbal signal of availability (as opposed to the message with no verbal signal) and we substitute these updated beliefs into Equation 2 to get the expected utility from calling the female parent after updating on the signal.

$$\begin{aligned}
\mathbb{E}(U_{if}) &= (1 - w_{if} - w_{im})\overline{q}_f \overline{r}_f + w_{if} \widetilde{qr}_{if}^{F}(x_{if}) + w_{im} \widetilde{qr}_{if}^{M}(x_{im}) + s_{if}^t - (\delta_{if} + c_i) && (7)\\
&= (1 - w_{if} - w_{im})\overline{q}_f \overline{r}_f + w_{if} \left[\lambda_f^F \overline{q}_f \overline{r}_f + (1-\lambda_f^F) x_{if}\right] + && (8)\\
&\quad w_{im} \left[\lambda_f^M \overline{q}_f \overline{r}_f + (1-\lambda_f^M) \rho_f x_{im}\right] + s_{if}^t - (\delta_{if} + c_i) && (9)\\
&= \overline{q}_f \overline{r}_f - (1-\lambda_f^F)\overline{q}_f \overline{r}_f w_{if} - (1-\lambda_f^M)\overline{q}_f \overline{r}_f w_{im} + && (10)\\
&\quad (1-\lambda_f^F) w_{if} x_{if} + (1-\lambda_f^M)\rho_f w_{im} x_{im} + s_{if}^t - (\delta_{if} + c_i) && (11)\\
&= \alpha_f + \eta_f^F w_{if} + \eta_f^M w_{im} + \gamma_f^F w_{if} x_{if} + \gamma_f^M w_{im} x_{im} + s_{if}^t + \varepsilon_{if} && (12)
\end{aligned}$$

where the last equation follows from the mapping below:

$$\begin{aligned}
\alpha_f &= \overline{q}_f \overline{r}_f - \overline{\delta}_f - c && (13)\\
\eta_f^F &= -(1-\lambda_f^F)\overline{q}_f \overline{r}_f && (14)\\
\eta_f^M &= -(1-\lambda_f^M)\overline{q}_f \overline{r}_f && (15)\\
\gamma_f^F &= (1-\lambda_f^F) && (16)\\
\gamma_f^M &= (1-\lambda_f^M)\rho_f && (17)\\
\varepsilon_{if} &= (c - c_i) + (\overline{\delta}_f - \delta_{if}). && (18)
\end{aligned}$$

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

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

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

Given the assumption that $\alpha_n = 0$, the other α_j intercepts are directly identified by comparing the proportions of decision makers who receive no signal and call parent j and the proportions who receive no signal and call neither parent. To separately identify γ_j^J and η_j^J, we need to create variation in the term $w_{ij}x_{ij}$, that is, the interaction of the indicator variable for whether a signal was sent (w_{ij}) and the value of the signal (x_{ij}). This variation must be distinct from the variation in w_{ij} alone. We achieve this by sending two values of the signal about each alternative j with known cardinal values. Specifically, we send both a positive signal and a negative signal about each parent and assume the values are 1 and -1.[2]

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

Lemma 1. *Given the assumptions of Sections 4.1–4.4, the reduced-form parameters α_j, γ_j^J, and η_j^J are identified for $j, J \in \{f, m\}$.*

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

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

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

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

Sending a signal ($w_{if} = 1$) with value $x_{if} = 1$ about alternative f and no signal about alternative m makes the deterministic part of utility for alternative f (i.e. Equation 12 without

[2]For a discussion of the impact of the chosen scale of signals, see Section G.6.

the error) $\alpha_f + \eta_f^F + \gamma_f^F \pm s_f^{h\boldsymbol{F}}$. We therefore have the following probabilities:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\frac{p_f^{n\boldsymbol{S},M}}{p_n^{n\boldsymbol{S},M}} = e^{\alpha_f - s_f^{n\boldsymbol{S}}} \Leftrightarrow \alpha_f - s_f^{n\boldsymbol{S}} = \ln p_f^{n\boldsymbol{S},M} - \ln p_n^{n\boldsymbol{S},M} \tag{20}$$

Adding Equation 19 from Equation 20 and then simplifying, we have

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

If we instead subtract Equation 20 from Equation 19 and then simplify, we have

$$\boxed{s_f^{n\boldsymbol{S}} = \tfrac{1}{2}\left[\ln p_f^{n\boldsymbol{S},F} - \ln p_n^{n\boldsymbol{S},F} - \ln p_f^{n\boldsymbol{S},M} + \ln p_n^{n\boldsymbol{S},M}\right]} \tag{22}$$

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

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

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

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

$$\frac{p_f^{l\boldsymbol{F},M}}{p_n^{l\boldsymbol{F},M}} = e^{\alpha_f + \eta_f^F - \gamma_f^F - s_f^{l\boldsymbol{F}}} \Leftrightarrow \alpha_f + \eta_f^F - \gamma_f^F - s_f^{l\boldsymbol{F}} = \ln p_f^{l\boldsymbol{F},M} - \ln p_n^{l\boldsymbol{F},M} \tag{26}$$

Subtracting Equation 24 from Equation 23 and then simplifying, we have

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

Likewise, subtracting Equation 26 from Equation 25 and then simplifying, we have

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

Now, if we subtract Equation (25) from Equation (23), we have

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

Substituting in for the reply to sender terms from (27) and (28), we have

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

Simplifying, we have

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

Combining Equations (21), (23), (27) and (29), we have

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

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

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

$$\boxed{\eta_f^F = \tfrac{1}{2}\left[\ln p_f^{h\boldsymbol{F},F} - \ln p_n^{h\boldsymbol{F},F} + \ln p_f^{l\boldsymbol{F},M} - \ln p_n^{l\boldsymbol{F},M} - \ln p_f^{n\boldsymbol{S},F} + \ln p_n^{n\boldsymbol{S},F} - \ln p_f^{n\boldsymbol{S},M} + \ln p_n^{n\boldsymbol{S},M}\right]} \tag{30}$$

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

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

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

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

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

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

G.3 Identification of Structural Parameters

Recall Equations (13), (14) and (16), which map the key reduced-form parameters for female parents to the key structural parameters for female parents:

$$\alpha_f = \bar{q}_f \bar{r}_f - \bar{\delta}_f - c \tag{13}$$

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

$$\gamma_f^F = 1 - \lambda_f^F. \tag{16}$$

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

Result 1. *Given the assumptions of Sections 4.1–4.4 and Lemma 1, the structural parameters* λ_f^J*,* λ_m^J*,* $\bar{q}_f \bar{r}_f$*,*$\bar{q}_m \bar{r}_m$*, and* $\bar{\delta}_m - \bar{\delta}_f$ *are identified for* $J \in \{f, m\}$*.*

Proof: γ_f^F directly identifies λ_f^F as $\lambda_f^F = 1 - \gamma_f^F$ in a simple rearrangement of Equation (16). Once we have λ_f^F, we combine it with Equation (14) to get $\bar{q}_f \bar{r}_f = -\frac{\eta_f^F}{\gamma_f^F}$. Finally, from Equation (13), we have $\bar{\delta}_f + c = -\frac{\eta_f^F}{\gamma_f^F} - \alpha_f$. Analogous equations for the male parent give us $\lambda_m^M = 1 - \gamma_m^M$, $\bar{q}_m \bar{r}_m = -\frac{\eta_m^M}{\gamma_m^M}$, and $\bar{\delta}_m + c = -\frac{\eta_m^M}{\gamma_m^M} - \alpha_m$.

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

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

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

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

Now rearrange and substitute in the beliefs to get

$$\alpha_f - \alpha_m = \overline{\delta}_m - \overline{\delta}_f + \overline{r}_f - \overline{r}_m. \tag{31}$$

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

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

G.4 Mapping Treatment Effects to Reduced-Form and Structural Parameters

If we include the treatment-specific reply-to-sender terms as the covariates in X_i, it is straightforward to map the coefficients from the treatment effects regression in Equation 1 to the reduced-form parameters from Equation 12. Both are displayed in Table A.3, where we use the no-verbal-signal treatment from the Baseline variation and the four signal treatments from the Equal Decision variation.

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

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

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

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

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

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

G.5 Model with decision maker characteristics

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

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

With decision-maker characteristics, Equation 2 becomes

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

Each type g of the decision maker makes their decision as in Section 4.3. 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 4.3 so that beliefs are independent across types of decision maker, e.g., that all $r_{ij,g} \sim \mathcal{N}(\bar{r}_{j,g}, \omega_j^2)$

are independent across g.

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

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

Equation (12) becomes

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

Similarly, equations (13)-(18) become

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

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

We then have the following identification result:

Result 2. *Given the assumptions of Sections 4.1–4.4 and this section, the reduced-form parameters $\alpha_{j,g}$, $\gamma^J_{j,g}$, $\eta^J_{j,g}$ and the structural parameters $\lambda^J_{f,g}$, $\lambda^J_{m,g}$, $\bar{q}_{f,g}\bar{r}_{f,g}$, $\bar{q}_{m,g}\bar{r}_{m,g}$, and $\bar{\delta}_{m,g} - \bar{\delta}_{f,g}$ are identified for $j \in \{f, m\}$, $J \in \{f, m\}$ and $g \in G$, G discrete.*

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

G.6 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 (i.e., η_j) does not change. However, although the signal's value is now assumed to be different, the term $(1-\lambda_j)w_{ij}x_{ij}$ in Equation 12 does not vary with our assumption about the value of x_{ij}. Instead, when we change x_{ij}, the value of $\gamma_j = (1-\lambda_j)$ adjusts to compensate since w_{ij} is simply an indicator for whether any signal is sent. Therefore γ_j is scaled in the opposite direction of the signal value. For instance, if the signals go from magnitude 1 to magnitude 2, γ_j is cut in half. The intercepts, α_j, do not change since they are entirely determined by the proportions of calls when there is no signal.

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

G.7 Risk Aversion

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

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

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

H External Validity

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. We fully acknowledge that gender identity takes more than two values, but we have started this research with the two ends of the gender spectrum (male and female).

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

We believe the direction of the effects of our high/low-availability messages would be the same for a variety of genders (e.g. two non-binary parents, same-sex couples), however we would expect No Signal inequality to be closer to zero in households with these gender identity sets. Nationally representative data indicates that same-sex households do not report wishing they were contacted more/less than they actually are by their child's school.[3]

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

We believe that any gender gaps that we document in our specific task in the school setting will generalize to other tasks in the school setting, such as picking up a sick child, or joining the Parent Teacher Association. Educators in our survey report that they would favor contacting the mother first in many of these scenarios (we discuss the survey in Section L.1). The gender distribution of these tasks is significantly skewed with mothers comprising almost 90% of Parent Teacher Association members and many surveys finding fathers self-report lower levels of involvement in their child's school activities, compared to mothers.[4]

I Efficiency

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

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

In our experimental data, even when the email comes from the father and he signals that he has high availability, 12% of the calls are still directed to mothers (Table 2). This indicates

[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.

[4]See our own survey in section L.3 and Daly and Groes (2017) https://archive.nytimes.com/parenting.blogs.nytimes.com/2009/01/06/dads-in-the-pta/, https://education.gov.scot/media/b3cn2mv5/nih327-dads-involvement-in-school.pdf

that households that want a more egalitarian division of child-related tasks and household labor, specifically fathers who want to be more involved, may be limited in achieving their goals in this area. Therefore, the current inequality in demands for parental involvement appears to be inefficient for some parents.

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

External decision-makers. Decision-makers may have multiple competing objectives. In our model (Section 4), the decision maker maximizing the likelihood of a useful response– a short-run outcome. However, in the long run, an entity (school, church, extracurricular program, doctor) may find it desirable to have a more diverse set of parents involved (e.g., not skewed toward mothers), and they may also prefer to have more parents (e.g., both parents versus one) involved (Clark et al., 1980). A less myopic decision maker may want to call the father even if they believe he is less likely to respond or may provide a less useful response. We believe work on these trade-offs is an important area for future research.

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

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

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.

J Ethics

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

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

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

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

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, and as such, participation in our study is arguably part of our subjects' regular job duties.

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

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

in Pager (2007):

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

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

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

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

K Data Collection and Matching

Emails and Phone Numbers To record phone metadata and voicemails we used a service called Callfire to set up a series of different phone numbers for our male and female parents. First, we set up a series of phone numbers with a generic voice mail box and text-message auto-reply saying that number did not receive text messages. We also set up email addresses with an auto-reply directing responders to please call instead of emailing. The exact email addresses that we sent our messages from were "erica@miller-family.net" and "roy@miller-family.net" for part of our data collection. We switched to emails from "audrey@the-johnsonfamily.net" and "curtis@the-johnsonfamily.net" for the bulk of data collection. We discuss the choice of exact names in detail below and in Section K.1. Due to constraints on email send limits, the follow-up emails sent after the first email which said the family no longer needed to talk were sent from "audrey@the-johnson-family.net " and "curtis@the-johnson-family.net."

Email is a common way for parents to contact schools. In our own survey, three-fourths of educators reported being contacted by parents via email at least once a month (Section L.1). These educators also reported that, when being emailed by both parents, a single parent emailing and cc'ing the other parent was more common than emails from a joint family email account. In one of our pilot data-collection efforts, we found that emailing from a joint email account lowered callback rates (Section K.1). Furthermore, we were concerned that a joint family email address might signal a more egalitarian family, which might bias our results towards finding more equal calls to mothers and fathers. As such, we decided to not use any joint family email accounts.

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

Messages We pretested our messages using a survey run on Amazon Mechanical Turk to select which messages gave the widest variation in self-reported likelihood of getting a call back. We also pretested our messages with a set of educators (see Section L.1) to ensure the messages seemed natural to this audience. Furthermore, we tested different versions of the two message variations we sent the most (Baseline and Equal Decision). The messages we sent were brief by design in effort to use less of the decision maker's time and to make our treatments about parent availability more salient. We did test longer versions of our two most-emailed messages, as detailed in Table K.1, but found that the difference in the callback rates was not statistically significant, nor was the proportion of calls to mothers versus fathers.

Table K.1: Longer Versions of Messages

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

K.1 Pilot Studies

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

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

Additionally, we tested two types of email accounts in our pilots, given that our survey of educators indicated that the use of a joint family email address was less common than the use of individual email addresses and cc'ing the other parent (Section L.1). We found that using a joint family email address (versus individual email addresses, with one parent cc'ing the other parent) decreased our callback rates by 9.2 percentage points ($p = 0.032$). With the evidence from both the pilot and the survey, we dropped the joint family email address in our full data-collection efforts.

K.2 Phone Call Data

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

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

July 2022 Phone Calls In July and August of 2022, we sent 72,136 emails. In the weeks following we received 30,214 calls. Much like our May data, these calls include spam calls. Our primary objectives with matching callbacks to specific schools 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.

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

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

L Survey Evidence

We collected data for this project via survey three times in 2022, 2023 and 2024. Here we describe those surveys in more detail. All the surveys were run on Prolific (IRB number STUDY00002608).

L.1 Educator Survey

In April 2022 we ran a survey of educators before we ran our main field study. People were eligible to take our survey if they were over 18, and reside in the US. We had 238 educator respondents in 2022. The goal of this survey was to check that the type of email we were sending to schools was appropriate. Over 50% of educators reported getting the most questions about school enrollment during the month of August. August was followed by the months of May, September, July, June and April (in that order) with about 18% to 28% of educators stating they got the most questions about enrollment in these months. About three-fourths of educators said that being contacted by parents was either very common (at least once a week) or somewhat common (at least once a month). When being emailed by both parents a single parent emailing and cc'ing the other parent was more common than emails from a joint family email account. Educators reported they contacted parents by phone about the same amount as they did via email, email being slightly more common.

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

We found that educators always reported a higher level of wanting to contact the mother instead of the father if they had to choose a single parent to contact about a child being sick (98% contact mom), volunteering at a book fair (96%) or career day (78%), school related payments (86%), or a child's allergies (97%). We allowed the educators to rank the following reasons for choosing to contact the person which were displayed in a random order: I expect this person to be more likely to respond quickly, I expect this person to be more likely to be 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.

L.2 Decison-Maker Survey

In April 2023 we ran a similar survey of adults who interact with parents, including educators. People were eligible to take our survey if they were over 18, reside in the US and regularly reach out to parents as part of their job. We had 377 respondents from a variety of persons who interact with parents (the most common were Teacher, Childcare provider, Medical Practitioner, Nurse, Sports Leader). Of the 377 respondents in 2023, 77 self-identified as interacting with parents in the role of "other." The primary purpose of this survey was to produce panel B of Figure 1. We randomly assigned whether a decision-maker was asked the following question about a [mother] or [father]: What proportion of the time do you contact the [father][mother] first if only contacting one parent first?

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

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

L.3 Household Surveys

In April 2024 we ran a survey of those who were based in the U.S., over 18, and identify as a parent in a two-parent household with both parents present. We had 349 respondents with 47% identifying as mothers and the rest identifying as fathers. One purpose of this survey was to measure how child related interruptions impacted mothers' and fathers' labor market and human capital decisions. We report our findings in Figure 6.

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

M American Time Use and the Current Population Surveys

Using ATUS data and following the methodology of Cubas et al. (2021), we replicate their finding that 35% of women experience a household interruption on a typical workday versus 20% of men (see Table M.1). In addition to the calculations we describe in Section 6, we can perform back-of-the-envelope calculations for the wage penalty associated with each type of the male-female split of interruption hours as well as each of the signals we send to decision-makers. We find that indeed our messages are effective at ameliorating the wage gap (reducing it from 7.7% to 7.1%, which is about a 10 percent reduction), but they do not (in this simple exercise) eliminate it even when 74% of interruption hours are pushed to the father by our Baseline variation with the Male High message. Thus, even if households strongly signaled father's availability (by designating him as a primary contact, for example), they would still not be able to achieve the desired split of external demands which would be associated with inequity in pay. Consistent with these findings, almost a third of parents in our survey report that the organizations are unable to honor their request to contact the father first *even* if the father does the majority of interacting with the organization.[6]

Table M.1: **Mean Hourly Wages and Household Care By Gender**

	(1) All	(2) Male	(3) Female
Ext: Incidence of household care 8 to 5	0.25	0.20	0.35
Int: Hours of household care 8 to 5	0.14	0.12	0.17
Log Hourly Wage	3.01	3.07	2.89
Hourly Wage	573	587	545
Observations	14386	9269	5117

Notes: The table is based on the work of Cubas et al. (2021). Respondents are 18–65 years old, who report usual weekly hours $\geq$ 35 in the CPS between 2003-2018, who are married with at least one child in the household, and whose diary day is a weekday. We also restrict the sample to those who report nonzero time spent on work-related activities at the work site during the diary day. Log hourly wage is constructed by dividing weekly earnings reported in the CPS by usual (total) hours worked last week. Weekly earnings that are top coded are recoded as 1.5 times the top-code value. The weekly earnings measure we use is reported only for wage and salary workers, so this table excludes self-employed workers. The regression also includes fixed effects for single years of age, detailed education categories, detailed race categories, and years. All regressions are weighted using ATUS weights.

Table M.2: **Correlations Between Log Hourly Wages and Household Care**

	Extensive Male Only Cubas et. al. Table 2 Col 1	Extensive Both Genders	Intensive Both Genders
Ext: Incidence of household care 8 to 5	-0.087***	-0.064***	
	(0.014)	(0.011)	
Int: Hours of household care 8 to 5			-0.034***
			(0.010)
Female		-0.283***	-0.289***
		(0.010)	(0.010)
Mean log hourly wages			
Male	3.023	3.023	3.023
Female	2.840	2.840	2.840
R^2	0.354	0.348	0.347
Observations	7937	12658	12658

Notes: Table notes the same as Table M.1.

[6] See Appendix L.3 for discussion of our household survey. Where when asked "Suppose that your partner does the majority of interacting with the organization (for example, drops the kids off and picks them up). What proportion of the time does the organization still contact you with a child-related inquiry?" women report being called 59% of the time, while men only 33%.