Intrahousehold Consumption Allocation and Demand for Agency: A Triple Experimental Investigation*

Uzma Afzal† Giovanna d'Adda‡ Marcel Fafchamps§ Farah Said¶

February 2021

Abstract

We conduct lab experiments to investigate demand for consumption agency in married couples from Pakistan. We find subjects are often no better at guessing their spouse's preferences than those of a stranger, and many subjects disregard what they believe or know about others' preferences when assigning them a consumption bundle. This suggests that individual executive agency within the household has instrumental value. We indeed find significant evidence of demand for agency in all experiments, and this demand varies with the cost and anticipated instrumental benefit of agency. But subjects often make choices incompatible with pure instrumental motives – e.g., paying for agency even when they know their partner assigned them their preferred choice. We also find female subjects to be quite willing to exert agency even though, based on survey responses, they have little executive agency within their household.

JEL Codes: D13, D91, D61, J16
Keywords: empowerment, gender equality, procedural equality, allocative efficiency, utility theory

*We thank Catia Batista, Erlend Berg, Doug Bernheim, James Berry, Emily Breza, Giacomo De Giorgi, Glenn Harrison, Muriel Niederle, Silvia Prina, John Quiggin, Kekti Sheth, Pedro Vicente, and Joseph Wang for their useful comments. We also benefited from comments from participants to the CIRPEE Workshop on Social Identity and Social Interactions in Economics (Quebec City, April 2016), the WZB "Field Days" Conference (Berlin, June 2016), the Advances with Field Experiments Conference (Chicago, September 2016), the CSAE Conference 2017 (Oxford, March 2017), the SEEDEC 2018 Conference (Wageningen, April 2018), the Third European Workshop on Household Finance (London, May 2018), the Applied Development Economics Conference (Lahore, September 2019), and the HKBU-NTU joint workshop (Hong Kong, December 2020), as well as from seminar participants at Stanford University, Georgia State University, the University of Oregon, University of California Merced, and Nova University in Lisbon. Funding for this study was provided by the International Growth Centre (IGC) and by the Economic and Social Research Council (ESRC).

†University of Nottingham
‡University of Milan
§Stanford University
¶Lahore School of Economics

1 Introduction

Much of the economic literature on intrahousehold consumption assumes allocative efficiency between spouses. In particular, the collective model of the household (Chiappori, 1997a,b; Browning and Chiappori, 1998) assumes that, subject to each member's weight in the decision process, the resulting allocations entail no waste of resources (Browning et al., 1994, 2010, 2013, 2014). Household members' bargaining weights are modeled to depend on exogenous distribution factors, such as labor market opportunities, gender ratios or divorce laws (McElroy, 1990; Lundberg and Pollak, 1996). The assumption of efficient allocations is based on the notion that spouses interact repeatedly, are perfectly informed about each other's preferences and outcomes, and are altruistic towards each other (Baland and Ziparo, 2017).

A large empirical literature documents a number of failures of allocative efficiency between spouses, particularly in developing countries. Spouses only partially share risk (Dercon and Krishnan, 2000); they allocate resources sub-optimally to production (Udry 1996) and consumption (Angelucci and Garlick, 2016; Rangel and Thomas, 2019); they hide their own income or savings from their spouses (Anderson and Baland, 2002); they lie (Ashraf et al., 2014); and incur costs to keep control over resources (Jakiela and Ozier, 2015; Ashraf, 2009). Moreover, domestic violence and the widespread practice of child and forced marriages suggest caution in assuming altruism and cooperation between spouses. Experimental evidence confirms that spouses fail to reap the gains from cooperation in the provision of public goods (Hoel, 2015; Barr et al., 2019).

We contribute to this literature by providing direct evidence on three rationales underlying the assumption of intrahousehold allocative efficiency. First, we test the presence of perfect information by examining spouses' knowledge of each other's preferences. Second, we evaluate the extent of altruism between spouses by investigating whether they respect the other's preferences when informed about them. Third, we ask whether efficiency losses can result from pure demand for decision agency, absent any instrumental value of decision-making power on outcomes.

We address these questions through an original experiment conducted with married couples from Punjab, Pakistan. We propose three variations of the basic experiment. In all versions, we ask participants to select between non-transferable and common consumption bundles; to guess what their partners would prefer to consume out of the available options; and to select a bundle for the partner to consume. Each participant can then decide whether to consume her own or the partner's choice. Experimental treatments vary whether an individual is informed of the partner's preferences before she makes a choice for her partner; whether an individual is informed of what her partner chose for her before selecting what to consume; and whether the partner within the experiment is a spouse or

stranger.

Our findings diverge from assumptions typically made in intrahousehold models. First, household members often do not know the preferences of their spouse. Second, even when they are informed about these preferences, they often choose to ignore them when deciding what the spouse will consume: we interpret this as a desire to make choices for others mixed with spite. Third, 11% of participants are willing to pay for agency even after they are informed that their partner selected their preferred item. This suggests that demand for agency is not solely instrumental, but that the decision process matters independently from its outcomes: some subjects value agency for its own sake.

These results apply to both men and women, but within the experiment women are more likely to exercise power and demand agency than men. In particular, information on the partner's preferences leads men, but not women, to revise their pick if it is revealed to be different from the partner's preferred item. Furthermore, women who discover to have guessed their partner's preference correctly are less likely to select their partner's pick. We suggest demand for agency among women as a potential explanation for these gender differences. Results from the third experiment, showing that women offered a microfinance product are less likely to demand agency and that their husbands are more likely to relinquish control over the decision, support this interpretation. We discuss other possible interpretations of our results, such as spiteful preferences, benevolent paternalism, and lack of understanding. We find limited empirical support for them.

The consumption bundles that we selected for the laboratory experiments are food and drinks. We focus on food and drinks for three main reasons. First, we could – and did– ask participants to consume them on the spot. Preventing transfers of the goods increases the meaningfulness of the choices made by subjects.[1] Second, decisions about what to eat or drink are made regularly within the household, often multiple times in the course of each day, by both men and women: men typically do most of the household purchases, and women do all the food preparation. Hence food and drink are a domain of choice over which both spouses have some experience of power and agency. For the experiment we select dishes and drinks that are part of the daily diet of the typical participant because we expect them to be associated with well-defined preferences and with better knowledge of the spouse's tastes over them. Third, food consumption is an important determinant of individual health and well-being. There are large and well-established economic literatures using differences in food consumption within the household as basis for intrahousehold welfare comparisons (Pitt et al., 1990;

[1] This is in contrast to the prevailing practice in laboratory experiments played with married couples, which typically use monetary payoffs (Jakiela and Ozier, 2015; Verschoor et al., 2019) or food items handed out in the lab but consumed at home (Bakhtiar, 2019): since these items can be transferred to others, decisions made in the laboratory can be undone outside of it, and the experimenter does not observe the effect of the treatment on consumption itself.

Behrman, 1993; Deaton and Paxson, 1998; Strauss and Thomas, 1998; Thomas, 1994; Hoddinott and Haddad, 1995; Haddad et al., 1997), and documenting differences in food consumption and nutritional level between husband and wife (Ferro-Luzzi et al., 1990; Dercon and Krishnan, 2000; Fafchamps et al., 2009; Brown et al., 2020). This makes food consumption an excellent vantage point from which to assess agency and welfare allocation between spouses.

To test the generalizability of our results to other decision domains and to alleviate concerns of lack of replicability of experimental results (Camerer et al., 2016; Christensen and Miguel, 2016; Andrews and Kasy, 2019), we run a third version of the experiment with higher-value stakes, in the form of personalized leather goods.[2] This experiment is embedded within a field experiment, described in Afzal et al. (2019), offering randomly selected women in our sample access to a microfinance product. The fact that we reproduce all the key features of our experiment – focus on agency and power, in-kind payoffs consumed in person, and samples of married couples – in three distinct studies sets our work apart in terms of breadth and depth.

Our study contributes to the large literature looking at the efficiency of household decisions and questioning the assumptions of allocative efficiency, perfect information and altruism between spouses. Many papers, relying on natural, field, and laboratory experiments, show that income pooling fails, a situation that implies inefficient consumption choices (Duflo and Udry, 2004; Castilla, 2015; Chen, 2013; Iversen et al., 2006; Kebede et al., 2014).[3] In the same vein, several experiments have varied the amount of information available to household members and found evidence of opportunism (Ambler, 2015; Castilla and Walker, 2013; Castilla, 2014; de Laat, 2014; Hoel, 2015). A number of authors have documented the negative efficiency consequences of mismatched preferences between spouses (Ashraf et al., 2014; Carlsson et al., 2009; Fiala, 2017; Schaner, 2012). Our contributions to this literature are to provide direct evidence of imperfect information on spouses' preferences over common goods; and to identify other possible sources of inefficiency, namely demand for pure agency and willingness to disregard others' preferences within households.

This paper is also related to a number of studies that have experimentally examined 'instrumental' willingness to pay for agency in games between spouses (Iversen et al., 2011; Mani, 2011; Jakiela and Ozier, 2015) and that have studied how inefficiency in household decisions varies with the asymmetry of control over resources between men and women (Deer and Twyman, 2012; Ashraf, 2009; Fiala, 2017) Bakhtiar et al. 2020). We contribute to this literature by offering a novel experimental design to

[2] To limit the scope for transferring the good to someone else, each leather good is inscribed with the name of the recipient, as illustrated in Online Appendix F.

[3] Other income pooling tests are provided by Munro et al. (2008); Bobonis (2009); de Brauw et al. (2014); Robinson (2012), among others.

investigate whether people value agency and why. Our approach is designed to distinguish different motives behind willingness to pay for agency, such as imperfect knowledge of others' preferences or unwillingness to respect them. This also links our study to the literature that looks at preference for agency in decision-making processes outside of the household. Existing evidence is consistent with our result that individuals value their decision power beyond its instrumental value (Güth and Weck-Hannemann, 1997; Fehr et al., 2013; Bartling et al., 2014; Owens et al., 2014) and prefer fair decision processes, over and above the fairness of the allocation (Bolton et al., 2005).[4]

The paper is organized as follows. Section 2 offers a conceptual framework to guide the analysis and interpretation of results. The details of our experimental design are presented in Section 3. Section 4 discusses empirical evidence from the two laboratory experiments. Additional evidence from the field experiment is presented in Section 5, while Section 6 concludes.

2 Conceptual framework

Economists typically characterize individual preferences as being defined over material consumption – possibly of self and others – which they write as a welfare function $W_i(c)$ where c is a consumption vector covering the joint and individual consumption of individuals i and j. In married couples, it is common to set $W_i(c) = U_i(c) + \alpha_i U_j(c)$ where $U_i(c)$ describes the preferences of individual i over his or her individual consumption, and $U_j(c)$ does the same for j, and α_i is a welfare weight – positive for altruism and negative for spite (e.g., Becker (1981), Browning and Chiappori (1998), Browning et al. (2014)). Other functional forms have been used to capture other-regarding preferences, such as inequality aversion (Fehr and Schmidt 1999) which combines altruism and spite. All these functional forms have one feature in common: the only arguments of the welfare function are material outcomes and how they affect individual utilities from consumption. The way by which these outcomes are achieved does not matter.

To illustrate what we mean by process preferences, we use the example of an everyday situation from which our experiment is inspired. Imagine that Jack and Jill are at the restaurant, ready to order. There are various combinations of dish and drink to choose from, all priced equally. Jack and Jill each have a well defined preference ordering of all combinations (c); this preference ordering is known imperfectly by the other. Jack picks for himself his favorite meal c^*_{Jack} and asks Jill what she wants. She picks c^*_{Jill} and Jack orders it. In this example Jack has full executive agency – he places the order – but he chooses to consult Jill since he does not know what she likes. Doing so does not detract

[4] A large literature on incentives and sanctions further demonstrates the presence of control aversion among some individuals (e.g., Falk and Kosfeld (2006)).

from his material utility since consumption choices are non-rival: Jack gains nothing, materially, by deviating from Jill's selection. Knowing this, Jill tells Jack her preferred choice and trusts him to order it. The outcome is efficient since they each get their preferred choice.

In this vignette, agents have standard utility functions and have no preferences over process. Could they demand executive agency? To investigate this, imagine that Jack has no altruism or spite towards Jill: $W_{Jack}(c) = U_{Jack}(c)$. Jack is thus indifferent whether Jill gets her preferred choice or not. It follows that if consulting Jill or remembering her order requires effort, Jack may, for purely instrumental reasons, choose to order at random for Jill. But if someone exogenously provides him with Jill's preferred choice and offers to bring it, there is no instrumental reason for Jack to deviate. Similarly, if Jack is told that Jill has already ordered, or will order on her own, there is no reason for Jack to object. In other words, since Jack has no instrumental reason to want to order for Jill, he is happy to delegate the decision to her. What he wants is executive agency over his own order. Since the contrast between these two types of executive agency pervades the entire paper, let us call *agency* an agent's control over their own consumption, and call *power* an agent's control over the consumption of another person. Furthermore, any of these demands is called *instrumental* if it serves to increase $W_i(c) = U_i(c) + \alpha_i U_j(c)$ through the consumption vector c. In our example with standard utility functions, Jack and Jill have a demand for agency for instrumental reasons, and they have no demand for power.

Now let's assume again that Jack orders Jill's preferred dish, which means that Jill's instrumental value of agency is 0, and let us introduce preferences over process. Imagine that Jill values agency over and beyond its instrumental value. As a result Jill prefers to order herself than letting Jack consult her and order for her. This can be represented as:

$$W_{Jill}(c^*_{Jill}, D^{Jill}_{Jack}) < W_{Jill}(c^*_{Jill}, D^{Jill}_{Jill}) \tag{1}$$

where D^j_i means agent i decides the meal of agent j. Here Jill's welfare is higher if she chooses herself even if she is certain to receive her favorite meal c^*_{Jill}. It follows that there exist a reduction on consumption ρ_{Jill} such that:

$$W_{Jill}(c^*_{Jill}, D^{Jill}_{Jack}) = W_{Jill}(c^*_{Jill} - \rho_{Jill}, D^{Jill}_{Jill}) \tag{2}$$

In other words, Jill is willing to pay for pure agency.

It is also conceivable that Jack prefers to order for Jill – he likes to be in control. This can be repre-

sented as:

$$W_{Jack}(c^*_{Jack}, D^{Jill}_{Jack}) > W_{Jack}(c^*_{Jack}, D^{Jill}_{Jill}) \tag{3}$$

Hence there exist an increase in consumption δ_{Jack} for which Jack is willing to delegate Jill's order to her. In equation (3) Jack derives no direct utility from what Jill consumes; he only enjoys ordering for her. In contrast, a spiteful Jack derives utility only if he chooses for Jill a consumption bundle c_{Jill} that is different from her preferred choice c^*_{Jill}:

$$W_{Jack}(c^*_{Jack}, D^{Jill}_{Jack} | c_{Jill} \neq c^*_{Jill}) > W_{Jack}(c^*_{Jack}, D^{Jill}_{Jack} | c_{Jill} = c^*_{Jill}) \tag{4}$$

In this scenario, if Jack is told c^*_{Jill}, he is *less* likely to select it – a direct manifestation of *spite*.

The reverse can also be true. For instance, Jill may dislike making the decision herself, she may prefer to delegate to Jack. This can occur for a variety of reasons, e.g.: Jill has internalized a social norm to this effect (e.g., Dean and Jayachandran (2019); Banerjee et al. (2018)); she fears retribution if violating this norm (e.g., Rao (1997); Bloch and Rao (2002); Pollak (2004)); she does not aspire to agency in general (e.g., Krishnan and Krutikova (2013); Alan et al. (2018); Bernard et al. (2016)); or she knows Jack likes to be in control and chooses to accommodate that. Whatever her reasons, we now have:

$$W_{Jill}(c^*_{Jill}, D^{Jill}_{Jack}) > W_{Jill}(c^*_{Jill}, D^{Jill}_{Jill}) \tag{5}$$

which, by the same reasoning as above, implies that Jill is willing to consume something different from her preferred consumption bundle, i.e., she is willing consume $c_{Jill} \neq c^*_{Jill}$, for c_{Jill} sufficiently close to c^*_{Jill}:

$$\exists\, c_{Jill} \text{ such that } W_{Jill}(c_{Jill}, D^{Jill}_{Jack}) = W_{Jill}(c^*_{Jill}, D^{Jill}_{Jill}) \tag{6}$$

By extension, even if Jill is unsure what Jack will choose for her (for instance because he does not know what she wants or does not care), it is still possible that Jill will let Jack decide for her:

$$\exists\, f(c_{Jill}) \text{ such that } E[W_{Jill}(c_{Jill}, D^{Jill}_{Jack})] \geq W_{Jill}(c^*_{Jill}, D^{Jill}_{Jill}) \tag{7}$$

where $f(c_{Jill})$ denotes the pdf of Jill's expected consumption when Jack chooses. The more c_{Jill} is likely to deviate from c^*_{Jill}, the less likely Jill is to delegate. For instance if Jill is matched with a pure stranger, the stranger is less likely to guess her preference, which reduces $E[W_{Jill}(c_{Jill}, D^{Jill}_{Jack})]$. The same holds if c_{Jill} is selected at random by a computer. Hence Jill is less likely to surrender agency when matched with a stranger or a computer. Similarly, if Jack is spiteful towards Jill, this also lowers $E[W_{Jill}(c_{Jill}, D^{Jill}_{Jack})]$ and reduces Jill's willingness to delegate. Hence we expect Jill to

demand agency when Jack has demonstrated a lack of altruistic feelings towards her – e.g., via a separately played dictator game.

3 Experimental design

In this section we describe the two laboratory experiments, and we turn to the field experiment in Section 5. The two laboratory experiments are modelled directly on the vignette above: couples are brought to the lab and they are invited to separately select a drink or a meal for themselves and for their experimental partner. Treatments are introduced that modulate information transmission and, at the end, subjects are asked whether they wish to reduce their consumption in order to ensure getting their preferred bundle. This allows us to estimate demand for agency and power by both spouses.

Formally, in both experiments each subject i is invited to select among several consumption bundles. Another person j is also invited to select a consumption bundle for i. Subject i is then offered a chance to pay something to ensure that receiving his/her selection. We expect i's willingness to pay for agency to be higher when the instrumental value of agency is larger – e.g., when j has little information about i's preferences and is less likely to respect those preferences. To generate variation in the instrumental value of agency, we first vary j's relationship with i: in one treatment, j is a spouse or close household member; in another j is a stranger. We expect a stranger to be less informed about i's preferences and less concerned about choosing the wrong bundle for i. Second, we vary the information that individual j is given about i's preferences. We expect an informed j to be more likely to accommodate i's preference when j cares about i – e.g., is a close relative or is otherwise altruistic. Hence the effect of information on accommodating i's stated preferences should be larger between close relatives than among strangers. The experiment is designed to test these predictions directly.

Concerning implementation, three priorities inform our choices in both experiments. First, we want to make the choice of consumption bundle final: as already discussed, the use of food and drinks to be consumed on the spot serves this purpose.[5] Second, we want to keep individual choices private, especially from their partners. This protects subjects from retribution and enables them to freely express their agency. To achieve this, we place male and female subjects in two different rooms for the whole duration of the experiment, i.e., until after they have consumed the food and drinks of their choice. Preventing communication distinguishes the experiment from how decisions are made within the family: typically spouses learn about each other's preferences by talking to each other. Communication, however, can also serve other purposes, such as imposing one's own views or will,

[5] In addition, it is frequent to offer refreshments to study participants and volunteers to community activities. This makes the experimental choice feel natural in our study setting.

and these are irreconcilable with our research questions. Third, we want the experiment to reveal whether spouses know and respect each other's preferences, even in the absence of communication: this requires choosing consumption goods that are extremely common among our study participants. In the first experiment, subjects select one of three common juice flavors – apple, pineapple and orange.[6] In the second experiment, participants select one of three staple dishes - two savoury and one sweet - that are cooked following standard recipes and are available from all type of food outlets - from street vendors to high-end restaurants. They also select one of three drinks: Coca Cola, water, or Rooh Afza, a local sweet drink.[7]

3.1 The first laboratory experiment

The first experiment (LAB1) was conducted in April 2015 in the city of Chakwal, Pakistan. There were 15 sessions with a total of 336 subjects. We invited female participants in a microfinance RCT and some of their randomly selected neighbors to come to a makeshift lab with their spouses. Those who could not bring their husband were told to come accompanied by the adult male with the greatest power over her household's finances.[8]

Upon arrival, men and women are seated in separate rooms. Each male and female subject is then matched with a partner in the other room. Half of the women are matched with the male household member they came with – Family Match treatment; the other half are matched at random with a male participants from another household – Stranger Match treatment. Subjects in the Stranger Match treatment are never told the identity of the person they were matched with.

Each subject is then invited to taste each of three possible fruit juices, upon which the subject ranks them by order of preference and selects a flavor as his/her preferred choice. Next, the subject is asked to guess the partner's ranking of the three flavors. Half of the subjects are then informed of the actual ranking of flavors by their partner and they are asked, with that knowledge, to pick a juice for the partner – but they are not told whether the partner receives similar information about them. The rest of the subjects are not told their partner's rankings and are asked to pick a flavor for the partner without having that information. Subjects are then individually told that a coin flip will determine whether they receives their own choice of juice or the juice picked for them by the partner. Before tossing

[6] These flavors are offered by every major juice brand, both local and international; and juice is consumed widely both by children and adults.

[7] The dishes were Chicken tikka boti, Biryani rice, and Zarda (sweet rice). While we do not have data on the share of Pakistani households familiar with these items, they all feature prominently in the Wikipedia page on Pakistani cuisine: `https://en.wikipedia.org/wiki/Pakistani_cuisine`

[8] About 74.4% of the participants came with their spouse; 14.9% with their son and 10.7% with another male household members (e.g., brother or brother-in-law).

the coin, however, subjects are offered the opportunity to instead consume half a glass of the juice of their own choice and forgo the coin toss. We also conduct three redistributive games to measure altruism: a dictator game; an ultimatum game; and a reverse dictator game. A detailed description of the experimental protocol is given in Online Appendix D1.

3.2 The second laboratory experiment

The second lab experiment (LAB2) was conducted in April 2017 in 15 different villages of the Faisalabad district. A total of 30 lab sessions were conducted with 358 married couples. The design and general testing strategy are the same as LAB1, but a number of modifications were introduced to verify the robustness of our initial findings and to introduce other sources of exogenous variation for identification purposes. The details of the experimental protocol are given in Online Appendix D2. The main modifications are the following:

1. The consumption bundle is expanded to include three food items and three beverages. In addition, the differences between the foods and beverages on offer are increased to raise the utility cost of not getting one's choice. The purpose of these changes is to increase statistical power.

2. All subjects play two rounds of the consumption game, one with their family partner and one with one of three stranger categories: Opposite Sex Stranger (OSS); Same Sex Stranger (SSS); or Computer Match (CM). The order of Family Match versus Stranger Match is randomized at the session level.[9]

3. Subject are asked to guess the food and drink rankings of their spouse and their randomly assigned partner. But, unlike in LAB1, guessing correctly is incentivized: those who guess the preferred food *and* drink of the spouse or partner receive PKR 100; those who guess only one of them receive PKR 50; others receive 0. Subjects are then asked to select a consumption bundle for their partner – except in the Computer Match treatment for which this decision is too artificial.[10] The computer does, however, select a random bundle for each of its assigned partners.

4. As in the first experiment, some subjects are informed of the preference rankings of their partner before choosing a bundle for the partner. Each subject assigned to the information treatment is told the partner's ranking in both rounds of the game. But subjects are not told whether the

[9] In half of the sessions, subjects start with the Family Match; the other half start with the Stranger Match. Within each session, one third of the subjects are assigned to OSS, one third to SSS, and one third to CM.

[10] Unlike in the first experiment, we no longer ask subjects to rank the available options and to separately pick a consumption bundle: the top ranked item is taken as the subject's choice. The purpose of this simplification is to make time for other features.

partner receives similar information about their own rankings. This aspect of the experiment does not apply to the Computer Match treatment.

5. The choice between a coin toss and half a glass is replaced by the following. Subjects are first told that they will receive the bundle selected for them by their partner. But they can forfeit a fraction of their show-up fee to get their own choice. Subjects are asked to make this decision for one of three possible forfeited amounts – 5PKR, 20PKR and 50PKR – equally randomized across subjects. Once subjects have made this choice, they are individually told what the partner has selected for them and they are asked again if they want to pay (e.g., 5 PKR, 20 PKR or 50 PKR) to get what they have selected for themselves.[11]

6. We use a dictator game to measure altruism towards each of the two partners – the Family Match and the Stranger Match. Participants matched with the computer only play the dictator game with respect to their Family Match.

Online Appendix C reports a discussion and analysis of statistical power. Detailed information about sample sizes in different experimental cells is reported in Tables C.1 and C.2. Power calculations for LAB1 and LAB2 are reported in Tables C.3 and C.4, respectively. Since all dependent variables of interest are dichotomous and all regressions are estimated using a linear probability model, estimated coefficients represent percentage changes. For this reason, we report all detectable effects in percentage points (pp). In LAB1, for one-way comparisons (e.g., informed vs non-informed; stranger match vs. family match; or male vs. female) the minimum detectable effect size is 12.5 pp; while for two-way comparisons (e.g., informed vs uninformed for women only), the minimum detectable effect size is 18.5 pp (Table C.3). These values are smaller in LAB2, given the larger sample size and the fact that we have two matches for each subject (Table C.4). For one-way comparisons, the minimum detectable effect size is 8 to 9 pp. This rises to 12 to 15 pp for two-way comparisons.[12] These effect sizes are in line with the magnitude of effects typically discussed in lab experiments. We believe that smaller effects, even if statistically significant, would not be of a magnitude sufficient to warrant much attention in a lab context. It follows that our lab experiments have sufficient power to be of interest.

[11] At the end of the session, a coin toss randomly selects one of the two rounds of the game each subject played. A second coin toss then decides whether the first agency or the second agency question of that round determines their consumption bundle. This ensures that both agency decisions in both games are fully incentivized.

[12] We also report minimum effect sizes for three-way comparisons, but these only apply to some heterogeneity analysis reported in Appendices.

4 Empirical results from the lab experiments

We now report on our analysis of LAB1 and LAB2. The field experiment is discussed in the next section. We articulate our discussion of the empirical results around the three main questions that our experiment is designed to answer. First, we aim to study how well household members know each other's preferences over common individual consumption goods. Second, we want to test whether they respect those preferences if they know about them, or whether instead they wish to exert power over each other's consumption. Third, we want to measure individual demand for agency, that is, whether individuals are willing to pay for the certainty of getting a consumption bundle of their choice. We isolate the instrumental motive from the pure agency demand by testing if subjects are willing to *pay* to get their own choice even if they may already be getting their preferred consumption bundle. Throughout the analysis, we check for gender differences in behavior.

4.1 Sample characteristics

Before presenting the empirical results, we briefly discuss sample characteristics. All our female subjects are adult women in the Pakistan Punjab. Descriptive statistics on the samples of the two laboratory experiments and of the field experiment are presented in Table 1. Most of the women are married and have children. The average female subject is 39 years old and about half are literate. Subjects belong to households with an average monthly expenditure of little less than Rs. 19000 ($190). In contrast to the LAB1 sample which consists of micro-enterprise loan borrowers, more than three-quarters of the LAB2 and field experiment samples are housewives. A little under 75% of all three samples consist of women who are married to the household head. Our data also include information on respondents' ability to make autonomous decisions and on whether women are consulted in household choices, in domains ranging from small consumption to large financial decisions. These data reveal a lack of executive agency among women, who generally report having to ask for permission from the spouse before making common household decisions. It also documents the limited importance attributed to women's opinion. In contrast, men rarely have to ask permission to make autonomous decisions, even about large financial expenses.[13]

These patterns are consistent with the broader gender inequalities characterizing South Asia, and Pakistan in particular. In these countries, women lag behind men across a wide range of domains, from education and health to labor market and economic opportunities, as well as political empowerment. The Global Gender Gap Index captures all these dimensions of inequality and ranks Pakistan 151 out of 153 listed countries (World Economic Forum, 2020). Pakistan's ranking is lowest in terms of

[13] Figures A1 and A2 show how executive and consultative agency within the household vary across decision domains.

economic opportunities – which include labor force participation, the earnings gap, and the presence of women in managerial positions – and in terms of health outcomes – proxied by life expectancy and the sex-ratio at birth. This is partly due to the widespread practice of *purdah* and restriction on female mobility (Field and Vyborny, 2016). About 85% of women report having experienced physical or sexual violence (OECD data 2017), as reported in World Economic Forum (2020) and indices of female access to justice are among the lowest in the region. These features make Pakistan a relevant setting in which to study demand for agency by women.

4.2 Knowledge of others' preferences

We examine the first research question: do subjects know the preferences of their partner? This is achieved by comparing a participant's actual juice (LAB1) or food and drinks (LAB2) rankings with those guessed by his/her partner. LAB2 offers a stronger test of this research question, thanks to the incentivization of correct guesses, which should deter subjects from purposely misreporting what they believe to be their partner's preferences. Online Appendix B provides descriptive statistics on the distribution of subjects' preferences and guesses in the two experiments.

We compare subjects' guesses to two benchmarks: random guesses and optimal guesses. Random guesses ascribe a probability for each ranking equal to the frequency of that ranking among subjects.[14] It represents a lower bound on guessing accuracy – subjects should be able to do better than that. When subjects are uninformed about the specific preferences of their assigned partner, the optimal guess is to pick the single most common rankings in the study population, perhaps differentiated by gender (Eckel, 2016). For instance, since 75% of subjects prefer soft drinks to other beverages, guessing 'soft drink' will be right 75% of the time when matched with an unknown subject. This assumes that subjects know what items are popular among other participants – but it does not require knowledge of all probabilities, only of the mode of each distribution. Subjects may approach this benchmark when they guess the rankings of a stranger. Finally, for those subjects matched with a computer in LAB2, we expect them to do no better than random guess.

Panels (a) and (b) of Figure 1 display the distribution of guessing accuracy by match type in LAB1 and LAB2, respectively. Within each Panel, bars show the share of subjects by number of correct guesses. In LAB1 (Panel (a)), subjects rank three items. Since guessing the top and bottom-ranked items automatically implies a guess about the third item, guessing accuracy only takes three values: 0, 1, or 2 correct guesses. In LAB2 (Panel (b)), subjects rank three food items and three drinks, so the

[14] Formally, the choice frequencies are $\{p_1, p_2, p_3\}$ with $\sum p_k = 1$: a random guess picks choice k with probability p_k.

maximum number of correct guesses is 4.[15] Each bar reports guessing accuracy for each matching treatment: Family and Stranger in LAB1; Family, OSS and SSS in LAB2. For comparison purposes, each Panel also includes the distribution of random guesses and that of optimal guesses.

In Figure 1 Panel (a), we see that LAB1 subjects do better than random guesses, but not by much.[16] This may be because the consumption items used in this experiment are not strongly differentiated. We also note that, when guessing the preferences of a stranger, subjects are 15.2 pp more likely to make no correct guesses than the optimal benchmark (two-sided t-test, p=0.001) – possibly because subjects are unsure about the modal preference. Some subjects do better than optimal guessing when matched with a spouse or family member in terms of guessing their full ranking, although not significantly so (p=0.194). This suggests that subjects have some information about their family members' specific preferences, but not perfect foresight: indeed, 45.2% of LAB1 subjects are unable to guess the top or bottom ranked item of their family member, significantly more than the corresponding share under optimal guessing (p=0.005).

In contrast, LAB2 guesses are much better than random (Figure 1 Panel (b)).[17] This is probably because guessing is incentivized and preferences are more contrasted than in LAB1: indeed, 74% of subjects prefer the same dish and 75% prefer the same beverage. Subjects do not outperform optimal guessing, however: only 11 to 17% of subjects manage to guess their partner's rankings correctly, compared to 20% with optimal guessing. The shortfall with respect to optimal guessing is larger whenever subjects are matched with a partner of the same sex.[18]

Using regression analysis to test the significance of these results, we see that subjects in the family matching treatment in LAB1, and in family and opposite sex stranger matching treatments in LAB2,

[15] Except for in the Computer Match where subjects only guess the computer's top ranked items. This mechanically limits the number of correct responses subjects can give to 2 instead of 4.

[16] The distribution of guess quality is significantly different between the random and both the Family and Stranger matching treatments (chi-square distribution tests, p=0.028 and p=0.081 respectively), and the share of perfect guesses is significantly higher in both the Family and Stranger treatments than in the random one (two-sided t-tests, p=0.007 and p=0.026 respectively). However, we see no statistically significant differences in the share of no correct guesses between the random and the Family (p=0.208), nor between the random and the Stranger (p=0.451) treatments. In what follows, we use t-tests when we report tests on differences between proportions, and chi-square tests when test differences between distributions.

[17] Indeed, when we compare the distribution of guess quality in the random benchmark with that in the Family, OSS and SSS matching treatments, we always find statistically significant differences (all p=0.000).

[18] Relative to the Optimal benchmark, the share of perfect guesses is significantly smaller under Family match (p=0.001) and SSS match (p=0.000). Pairwise chi-square distribution tests between the Optimal and the Family, OSS and SSS treatments all yield significant differences (p=0.0000, p=0.001 and p=0.002 respectively). Only in the Family match treatment the share of no correct guesses is significantly smaller than in the Optimal benchmark (p=0.073).

do significantly better than random guessing, and not significantly worse than optimal guessing.[19]

Taken together, these results indicate that individuals in our sample population know relatively little about the specific preferences of their family members: they are not better at guessing the preference rankings of a family member than they are at guessing that of a stranger of the opposite sex: in both Panels (a) and (b) of Figure 1, the distribution of guess accuracy for family members lies virtually on top of that for strangers of the opposite sex.[20] This implies that agency over own consumption may have instrumental value, particularly for individuals who do not have the modal preference. There are no significant differences in guessing ability between genders.

4.3 Respect for others' preferences

We now turn to the second research question: do subjects respect the preferences of their partner? We investigate this in two ways: (1) we test whether uninformed subjects pick for their partner what they *guess* is their partner's preferred option; and (2) we test whether informed subjects pick for their partner what they have been *told* is their partner's preferred option. We also break down the marginal effect of information depending on whether the subject initially guessed correctly or not. We take *not* respecting the partner's preferences as a sign of demand for power, i.e., for exerting control over others' consumption. Alternative interpretations are possible, which we discuss in Section 4.5.

Overall, a large share of subjects disregard what they believe or know about their partner's preferences when assigning them a consumption bundle. In LAB1 on average, only 62.3% of subjects pick for their partner what they think is the partner's preferred item (uninformed treatment), and a lower 51.8% pick for their partner what they know is the partner's preferred item (informed treatment). We observe similar shares in LAB2: uninformed subjects pick for the partner what they believe to be the partner's

[19] Regression results for both experiments are displayed in Appendix A Table A1. A disaggregated analysis of LAB2 data is reported in Online Appendix B. A back-of-the-envelope calculation based on the proportions reported in Panel 5 of Online Appendix Table B.1 indicates that about 43% of subjects guess the modal (i.e., optimal) choice for their partner, around 25% pick their own preferences and the rest randomize equally among all three options. Since we find no correlation in preferences between family members, this pattern of behavior is consistent with the findings reported here using a more succinct methodology.

[20] In LAB1, the distribution of guess quality is not significantly different between the Family and the Stranger treatments, (p=0.224), nor is the share of perfect and no correct guesses (p=0.703 and p=0.663 respectively. Disaggregated analysis of LAB1 data available in Figure B.1 in the Online Appendix). In LAB2, we similarly observe no statistically significant differences between the Family and OSS treatments (chi-square p=0.517; share of no correct guesses p=0.927; share of perfect guesses p=0.267); nor between the Family and SSS (chi-square p=0.022; share of no correct guesses p=0.813; share of perfect guesses p=0.167). Disaggregated analysis of LAB2 data, shown in the Online Appendix B Figure B.2, confirms that this holds separately for drinks or food items, and for most or least preferred items. The only slight exception is that subjects are more likely to guess the food preference of a family member when it is the same as their own – perhaps due to a salience effect. We also find that the food and drink preferences reported by lab subjects are not more correlated between family members than they are between randomly matched strangers.

preferred food and drink 66.2% and 69.3% of the time, respectively. When informed, the proportion of 'correct' choices for the partner decreases to 57.7% for food and 61.9% for drink.

We regress respect for the partner's preference on gender-matching treatment dummies, controlling for whether the item is a drink, the amount given by the subject to his/her partner in the dictator game, and the amount received by the subject from his/her partner in the same game. We regard the amount given as proxying for altruism towards the partner, and the amount received from the partner as possible proxy for the altruism that the subject expects from the partner.[21] Males matched with a family member are the default category.[22]

Table 2 presents regression results for uninformed participants in LAB1 (columns 1 and 2) and LAB2 (columns 3 and 4). The dependent variable is a dummy equal to 1 if the subject picks for the partner what she guessed the partner prefers. For LAB1, we see from the intercept in column 1 that 74% of male subjects matched with a family member pick what they guess their partner prefers. This proportion falls by 13 percentage points for males matched with a female stranger, but the difference is not statistically significant. In contrast, women are on average 19 percentage points *less* likely than men to select what they guess is their family member's preferred juice, and this difference is statistically significant at the 5% level. A similar – though less precisely estimated – coefficient is found for female subjects matched to a stranger. Results for LAB2 are very similar: 83% of uninformed male subjects pick the item that they guess their family member prefers, a proportion that falls by 10% when the wife is the decision maker – a difference that is statistically significant at the 1% level. Results are robust to controlling for own and partner's altruism (columns 2 and 4). Taken together this evidence suggests significantly more demand for power among female than male participants, particularly towards household members.

In columns 5 – 8 of Table 2 we report coefficient estimates for informed participants. Here the dependent variable equals 1 if the subject picks for the partner what the partner ranked highest. For LAB1 subjects, we find that 61% of male subjects accommodate the preference of a female family member, but only 36% of them accommodate the preference of a female stranger – a difference of 25 percentage points, significant at the 5% level (column 5). The corresponding frequency for female subjects matched with a family member is a lower 56%, again suggesting that they are less likely to respect the preferences of their male partner. But women are more likely to respect the preferences of a male stranger. The differences between male and female subjects are not statistically significant,

[21] Since subjects only discover what they receive from their partner at the end of the experiment, our measure is likely to contain measurement error. Omitting the two controls for altruism does not change the results.

[22] The minimum detectable effect sizes are 12.5 and 8-9 percentage points for LAB1 and LAB2, respectively (Tables C.3 and C.4 in the Online Appendix).

however. A similar pattern is observed for LAB2 subjects in column 7: 74% of informed men pick their spouse's preferred food compared to 62% of women – a difference that is statistically significant at the 1% level. We also find that male subjects pay less respect to the preferences of strangers – significantly so for strangers of the opposite sex. These results on informed subjects confirm that both male and, to a greater extent, female subjects demand power over others.

Next, we examine the effect of information depending on whether the subject initially guessed the preferred item of the partner or not. We have seen that many uninformed participants pick for their partner a bundle *different* from what they believe their partner wants. Informing these subjects of their partner's true preferences could induce them to revise their selection *away* from what the partner likes, i.e., when informed they would be *less* likely to pick what their partner wants.[23]

To investigate these possibilities, we estimate the effect of the information treatment on the probability that the subject picks her partner's preferred bundle, *conditional* on whether the subject guessed the partner's top choice or not. Results are presented in Appendix A Figure A3 and Table A2. We find similar results in LAB1 and LAB2. Among subjects who guessed wrong, information increases the probability of picking the partner's choice. But the increase is only significant among male subjects matched with a family member: as a result of being informed of their partner's preferences, male subjects are 27 pp (LAB2) to 56 pp (LAB1) more likely to choose their partner's preferred option (both $p<0.001$). Among female subjects who guessed wrong, the effect of information on picking the correct bundle for a family member is small in magnitude – between 4.2 and 4.5 pp – and not statistically significant, and it is even smaller when matched with a stranger, suggesting that women who guessed wrong do not adjust their choice once informed of their mistake. This is also true of males subjects matched with a female stranger. The effect of information is more contrasted among subjects who guessed right: interestingly, both in LAB1 and LAB2, information does not increase the probability of selecting the partner's preferred item. Female subjects matched with a family member, in particular, are less likely to pick the right juice if they see their initial guess confirmed in the information treatment, although this effect is not statistically significant.[24]

Taken together, these results indicate that male and female participants behave in different ways: female participants are less likely to respect the top choice of a male partner, suggesting a demand for power. We do not observe a similar behavior among men, at least within the confines of our food-and-drink experiments. We also find that subjects who guessed wrong are more likely to revise their

[23] We do not anticipate a similar effect among subjects who correctly guessed their partner's preferences, except perhaps for a positive confirmation effect: providing information confirms what the partner selected and comforts the subject in their choice.

[24] The minimum detectable effect size is around 18 pp for LAB1 and 12 to 15 pp in LAB2 – see Tables C.3 and C.4.

choice when informed of their partner's preferences. But subjects who guessed right are equally likely to pick wrong when informed. This suggests that picking the wrong choice is not due to uncertainty about preferences, it is a deliberate decision to override the partner's choice.

4.4 Demand for agency

We now tackle the third research question: do subjects have a demand for agency over their own consumption? We define demand for agency as a dummy equal to 1 if a subject refuses the coin toss and prefers to receive half a glass of her preferred juice in LAB1; and as a dummy equal to 1 if a subject prefers to incur a monetary reduction of the show-up fee than receiving the food and drink that the partner has selected for her in LAB2.

We first investigate demand for agency by observing what proportion of subjects exercise it in the two laboratory experiments. Second, we ask whether the demand for agency increases with its instrumental value, that is, with the risk of not getting one's choice. To do this, we predict from available data the probability that the subject's ranking is respected by her partner using the various treatments – e.g., Stranger Match, Informed treatment – as predictors and we test whether there is more demand for agency when the predicted probability of getting the preferred pick is low. Third, we exploit a specific feature of LAB2 to test whether subjects have a demand for *pure* agency. This test is implemented by investigating what proportion of subjects are willing to pay to receive their selected bundle even when the food or drink that their partner selected for them is identical to what they selected for themselves. Throughout the analysis, we examine gender differences in demand for agency.

4.4.1 Overall demand for agency

The share of subjects exerting agency by gender and matching type is shown in Figure 2 Panels (a) and (b) for the LAB1 and LAB2 experiments, respectively.[25] The Figures pool the informed and uninformed treatments. A sizeable proportion of subjects in both experiments are willing to pay for agency: 23% in LAB1 and 26% in LAB2. There are, however, differences in pattern by gender and matching type across the two experiments: women in LAB1 are slightly less likely than men to demand agency when matched with a family member, but much less likely to do so when matched with a stranger (Figure 2, Panel (a)). The latter difference is statistically significant ($p = 0.048$). Other pairwise differences are not.[26] It is unlikely that the fear of reprisal explains why women are less willing to exert agency when matched with an anonymous stranger. Indeed, by experimental design,

[25] These results are also presented in regression format in columns 1 and 3 of Table 3.
[26] The minimum detectable effect size is approximately 12.5 pp in LAB1 – see Table C.3.

the identity of the stranger – who sits in another room – is never revealed and the risk of retribution is thus negligible. It is more likely that women have more demand for agency when matched with their husband or male guardian. We check the robustness of this finding below.

In contrast, LAB2 subjects all demand agency more often when matched with a stranger, whom subjects may expect to be less familiar with their specific preferences (Figure 2 Panel (b)). For men, the magnitude of the difference is fairly small, but it is statistically significant relative to family matching (p = 0.026 for Family vs. OSS; p = .057 for Family vs. SSS).[27] The difference is larger for women, 40% of whom prefer to forfeit part of their show-up fee to receive the food and drink of their choice rather than something selected by a stranger or a computer. We also note a large significant difference in willingness to pay for agency between male and female LAB2 subjects, across all matching types: female subjects are on average 56% more likely to pay for agency than men (p = 0.000). Finally, subjects of both genders are less likely to pay for agency when matched with a stranger of the same sex instead of the opposite sex, but the difference remains largest for women.

LAB2 subjects are given the opportunity to reject the random choice made by a computer. The reason for including this treatment is the concern that subjects may be reluctant to challenge the choice made by another subject, e.g., out of courtesy or for fear of offending them. If this is an issue, we expect more rejection of computer choices since they do not involve human intervention and the computer cannot possibly be 'offended' by rejection. We find no evidence of such concerns: the subjects' propensity to reject the choice of the computer is not larger than that of a stranger of the opposite sex (p = 0.380 for men and p = 0.716 for women). This offers reassurance that observed willingness to pay for agency is not self-censored.

To recap, we find that female subjects often – though not always – display a higher willingness to pay for agency than men. This evidence is consistent with our earlier results on higher demand for power among female subjects.

4.4.2 Instrumental value of agency

Next we investigate whether demand for agency varies with its instrumental value. To this effect, we test whether subjects are more willing to pay for agency when the expected gain is larger. As we have seen, subjects are often unaware of their partner's preference rankings – and even when they are aware, they often choose to ignore them. Having executive agency over own consumption can thus increase the likelihood of receiving one's preferred bundle. The question is whether subjects are more

[27] The minimum detectable effect size is approximately 8-9 pp in LAB2 – see Table C.4.

willing to pay when the chance of getting their preferred bundle is lower.

To test this hypothesis, we construct a variable proxying for the expected instrumental value of ex ante agency. How this proxy is constructed is described in detail in Appendix B, but the logic is as follows. We start by using the choices made by partners in the experiment to predict what a rational expectation subject would expect to receive. The difference between this expectation and what the subject would choose for herself is our proxy for the ex ante instrumental value of agency. We then regress ex ante demand for agency on this proxy variable – one for food and one for drinks in LAB2 – and on the cost of exerting agency, which is varied randomly across subjects in LAB2. We control throughout for gender-matching treatment. Estimation results are presented in columns 2 and 4 of Table 3. We see that in LAB2 subjects are significantly more likely to pay for ex ante agency when its instrumental value is high for the choice of food. In both LAB1 and LAB2 the instrumental value of the choice of drink is not statistically significant. The cost of agency has the expected sign and is borderline significant. These findings are consistent with the idea that demand for agency is at least partly driven by instrumental considerations.

We nonetheless find that gender and matching dummies remain strong predictors of ex ante agency, suggesting the presence of systematic variation in demand for agency that is not driven by its instrumental value. Comparing columns 1 and 2 for LAB1, and columns 3 and 4 for LAB2, we see that the coefficient on the matching dummies are robust to the inclusion of controls for the instrumental value of agency. In particular, focusing on LAB2, we observe that men matched with strangers are more likely to exert agency than men matched with their spouse or family member. The results also confirm our earlier finding from Figure 2, Panel (b) on LAB2, namely, that women are more likely to exert agency than men in all stranger matching types.

4.4.3 Demand for pure agency

To investigate the fundamental nature of demand for agency, we use data collected in LAB2 only. In that experiment, we first ask each subject whether she wishes to pay to receive the food and drink she has selected rather than what her partner has selected for her. We then tell the subject what her partner has actually selected for her and we ask her to confirm if she wishes to pay to have what she herself selected. We refer to answers to the second question as *ex post* demand for agency, to distinguish it from ex ante agency – captured by answers to the first question and discussed above. If demand for agency is driven solely by instrumental considerations, we should observe no demand for agency when the partner's choice matches the subject's. On the other hand, when the partner's choice is revealed to be different, we should observe an increase in ex post demand for agency relative to Figure 2. In contrast, if subjects demand agency for its own sake, they may be willing to pay to get

what they selected, even if it is the same as what their partner selected for them – as in equation (1).

Panel A of Table 4 reports the proportion of subjects who demand ex post agency, overall and by gender, while Panel B distinguishes between subjects who exerted agency ex ante or not.[28] What we observe does not suggest that demand for ex post agency is solely driven by instrumental considerations. Among subjects whose partner selected the correct food and drink, 11% nonetheless opt to pay for agency. At the same time, 83% of subjects (100% - 17% – see column 1 row 1) refrain from exercising agency even though they *know* that the partner has not selected what they prefer. In fact, only 53% of subjects who demanded agency ex ante still demand it ex post after discovering that their partner has *not* selected their preferred food and drink. This suggests reluctance in explicitly defying a specific decision made by someone else. We also find that, among those subjects who did not demand agency ex ante, only 4% of those who discover their partner got them the wrong food and drink decide to pay a small fee to get their preferred bundle. Furthermore, about the same proportion of subjects insist on exerting agency even when the partner gets their food or drink right. On the other hand, among those subjects who exerted agency ex ante, we find that 36% of them continue to exert agency ex post even when they are getting the right food and drink from their partner. Results are shown in panel B of Table 4. These findings are difficult to reconcile with a pure instrumental value of agency: there is too much demand for agency when the partner's selection is correct, and not enough when the partner is wrong.

It remains that ex post demand for agency falls when the partner gets the correct food. To test this formally, we regress demand for ex post agency on the actual value of agency – which is known by construction.[29] The results, shown in Table 5, are not too dissimilar from those shown in Table 3: ex post demand for agency increases with the gain from agency and falls with the price of agency. But estimated coefficients are smaller than what pure instrumental motivations would imply.[30] We also find that women are less likely than men to exert agency ex post when matched with a stranger, hinting at a reluctance in going against that person's decision. This suggests that at least part of the ex ante agency exercised by women in our experiment reflects demand for agency within the household, and that much of it is driven by non-instrumental considerations.

As a result of all the behavioral patterns documented here, a large proportion of subjects do not

[28] In Panel A the minimum detectable effect size is 8-9 pp; it is approximately 12 to 15 pp in Panel B – see Table C.4.

[29] The gain from agency is the difference between the partner's pick for the subject and its ranking by the subject; the price of agency is varied by experimental design.

[30] It is possible for a female subject to expect to have to relinquish all her lab earnings to her husband after the experiment is over. In this case, if she has any demand for agency, she would be willing to pay the maximum 'cost' for exercising it. As a result, the dollar 'cost' of agency may not have much explanatory power. This may explain why we see a significant but not a large coefficient on this variable in Table 5.

receive their preferred bundle. Part of this is due to the fact that their partner either does not know their preferences or chooses to ignore them. Moreover, it may be that subjects are not particularly good at targeting agency to situations where they are least likely to receive their preferred choice. For instance, in LAB2, ex ante agency increases the proportion of subjects who receive their preferred food and drink only from 48% to 62%. More surprising, ex post agency reduces this proportion to 56% – primarily because many women who chose to exert agency ex ante refrain from doing so ex post, even though their partner selected the wrong bundle.[31] This is quite remarkable given that, in the context of the experiment, individual consumption choices are deeply shrouded from their partners. This suggests that attitudes towards agency are internalized by subjects and at least partly enforced through internalized roles and norms. Some women do, however, demonstrate a demand for pure agency.

4.5 Discussion

Contrary to what is regularly assumed in much intrahousehold analysis, we have found evidence that: spouses are often ill-informed about each other's consumption preferences; they display a high propensity to disregard what they believe or know about each other's preferences; and their demand for agency is only partly explained by instrumental motives. We also find significant gender differences in demand for control over others' consumption and own consumption. We interpret this evidence as suggesting demand for power and agency, but other interpretations are possible, and we discuss each one in turn in what follows.

4.5.1 Spite

First, our theoretical model allows for the possibility that not respecting others' preferences generates direct positive utility, i.e., that spouses hold spiteful preferences – see equation (4). We expect spite to be more frequent in less cooperative households. We test this implication by adding to the regressions Table 2 various proxies of intrahousehold cooperativeness constructed from in the data: spouses' age difference; the woman's ability to take household decisions without permission; and answers to hypothetical questions on whether the husband would respect the woman's preferences if made aware of them. In addition, we estimate the regression in Table 2 separately for couples with above or below-median age difference, female decision-making autonomy, woman's age – which we use as a proxy of length of marriage – and similar or different tastes – proxied by whether the man and the woman

[31] Without exercising agency, 52% of LAB2 female subjects would receive their preferred food and drink. This proportion rises to 68% when ex ante agency is taken into account. But it falls back to 59% with ex post agency instead. The corresponding proportions for men are 43%, 55% and 53% – confirming that reluctance towards ex post agency is virtually non-existent among male subjects.

have the same top-ranked item. We expect more harmony, on average, in couples who have managed to remain married longer, have a smaller difference in age (a large difference is correlated with child marriage), grant more agency to women, and share similar tastes.

We find that none of the proxies of intrahousehold cooperativeness has a consistent statistically significant effect on the likelihood of choosing, for the partner, what is believed or known to be her preferred item. Importantly, our main results on differences across gender-matching pairs are robust to the inclusion of additional controls, and across sub-samples.[32] The analysis is reported in Online Appendix A.1.

4.5.2 Food choices and paternalism

Second, it is conceivable that our results are specific to the food choice domain: women may be accustomed to decide what people eat at home. Consequently, they may regard selecting food for others – and deciding which foods are healthy for others to consume – as their prerogative. Similarly, women in the study population may feel entitled to reject food selections imposed by men because they are used to exert agency on food preparation at home. While we cannot entirely rule out these possibilities, we offer three pieces of evidence that speak to their relevance.

First, we perform a simple check on the possible presence of a paternalistic motive. In South Asia, as in many other countries, many people perceive sugary drinks as unhealthy. Experimental subjects may thus override the preferences of others when it stops them from consuming sugary drinks. To investigate this possibility, we re-estimate the LAB2 portion of Table 2 for drinks only, to see whether this process can explain our results. We find no evidence in this direction.[33] Second, we exploit the fact that, in LAB1, the family partner was the female subject's son in 14.9% of the subject pairs. Paternalistic motives should be stronger for women matched with their sons.[34] Yet we find no statistically significant difference in results for this sub-sample.

Third, we take advantage of a feature specific to the LAB1 experiment, namely, that all subjects

[32] When we run similar analysis on the other outcome variables, we find that individuals in more cohesive couples appear to be better at guessing their partner's preferences. Measures of intrahousehold cooperation do not significantly affect our findings on demand for agency, however. This latter finding is inconsistent with our interpretation of gender differences in demand for agency as indicating pent-up demand by women, a point which we discuss further in Section 5. Results are shown in Online Appendix A, in Tables A1 - A3. Results from the subsample analysis are shown in Online Appendix A Tables A4 - A7.

[33] Soft drinks are the most popular drink, chosen by 75% of subjects. Rooh-afza is a local sugary drink chosen by 11.5% of subjects. The only non-sugary drink is water, which is selected by 13.5% of subjects. We find no evidence that subjects behave differently when their partner selects water instead of a soft drink.

[34] Table A3 in Appendix A provides the results. We cannot conduct a similar exercise for LAB2 since all participants in LAB2 were married couples.

play an ultimatum game against their assigned partner. Since rejection in the ultimatum game results in a monetary loss equal to the offer made, we use the propensity to reject a partner's offer as an alternative measure of costly demand for agency.[35] Results are shown in Online Appendix A.2 Table A8 and illustrated in Figure A4. All subjects are less likely to challenge an offer made by a stranger. Women are more likely to reject at least one of the offers made, irrespective of matching type: 36% and 86% (29% and 78%) of men and women, respectively, reject at least one hypothetical offer made by an accompanying family member (by a stranger). Women make significantly lower offers than men – an average of PKR 512 compared to PR 572 by men (p = 0.018). Overall, both male and female subjects tend to accept the offers made, indicating that they collectively see rejection as a costly action they can ill afford. Yet women are much more likely than men to reject at least one offer, an action that is consistent with a desire to assert agency over the outcome of the game without risking losing too much. This is reminiscent of the distinction we observe between the exercise of ex ante and ex post agency among women. To sum up, the evidence from the ultimatum game suggests that female demand for agency goes beyond protecting a small reserved domain of autonomy from husband interference.

4.5.3 Lack of understanding

Finally, we examine whether our results could be due to confusion. First, we exclude from the LAB1 analysis those subjects who did not pick for themselves their own top-ranked item – a choice that might signal lack of understanding. We find that our LAB1 results are robust across sub-samples (see Table A4 in Appendix A).

Second, if lack of understanding accounts for our results, we should observe learning over time. To investigate this possibility, we exploit the fact that in LAB2 subjects played two experimental rounds, one with a family match and one with another partner. Since the order of the two rounds is randomized, we expect order effects to be relevant only if learning takes place over the course of the experiment. Instead we find that order-effects are never statistically significant (see Table A5 in Appendix A).

[35] Since the game is played using a strategy method, we know each subject's response to all possible offers – and we do not have to worry about endogeneity of rejection with respect to the amount offered. The strategy method raises concerns of confusion among subjects. We assessed participants' understanding of the game through three hypothetical questions. The great majority of subjects answered all questions correctly. To err on the side of caution, we only use the 318 subjects who demonstrated a clear understanding of the ultimatum game, but similar results are obtained if we use all subjects.

5 The field experiment

In addition to the two lab experiments we have discussed, we implemented a simplified version of our design as a field experiment to verify the external validity of our lab findings. The detailed experimental protocol is presented in the Online Appendix D3.

5.1 Setting and design

Subjects are clients of a rural microfinance institution participating in a randomized controlled trial of a microfinance product with credit and savings characteristics (see Afzal et al. (2018) and Afzal et al. (2019) for details). The subject population is therefore similar to that of laboratory experiments 1 and 2, i.e., primarily middle-aged women and their husband. The RCT spans a period of one year, with a baseline survey at the beginning and an endline survey at the end. The experiment was conducted with 1991 respondents, 80% of whom were married.

In the baseline survey we tell participants that, as a token of our appreciation for their participation in the RCT, at endline we would offer them a small monetary reward plus one of three possible leather items engraved with initials of their choice. The three items to choose from were: a woman's purse; a man's wallet; and a child pencil case. A photograph is provided in Online Appendix F. The items were chosen to be relatively similar in quality and value, but to be desirable and gendered – i.e., for an adult man or woman – so as to induce potential competition between spouses.[36] Initials were added to make the item individual-specific in order to minimize the resale and gifting of the good to someone else.

At baseline, a female enumerator asked each female participant to choose one of these three objects and to tell us which initials to inscribe on it – e.g., her initials on the purse, her husband's initials on the wallet, or her child's initials on the pencil case. The same question was separately asked to the husband by another enumerator. One year later at endline, we revisit both husband and wife and offer them the following choices. We remind the husband of the item he had chosen at baseline, and ask him to confirm or revise his choice. We tell him that a coin toss would decide whether the household receives his preferred item or his wife's preferred item. We then ask him whether he wishes to delegate the final selection of the item to his wife. To those who refused, we ask whether they are willing to delegate the choice to their wife in exchange for a monetary transfer. Two price levels (50 PKR and 200 PKR) are offered to those who refused.

We separately tell each wife that a coin toss would determine whether we would offer her household

[36] The pencil case can be regarded as a household public good from the point of view of the spouses.

the item she had chosen at baseline, or the item her husband had just selected.[37] We tell the wife that she could secure her chosen item by forfeiting a fraction of the monetary compensation for participating in the study. Two separate price levels are offered, and responses are recorded for both. These choices are offered to each female participant even if her husband decided to delegate the choice to her. We then reveal whether her husband has delegated the choice to her, in which case the female participant receives what she selected and forfeits nothing. If the husband has not delegated, we first select one of the two price levels using a coin toss. If the coin toss yields a decision not to pay for executive agency, we proceed to a second coin toss to determine whether the offered item is her choice or her husband's. Otherwise the female participant forfeits the selected amount from her participation fee, and receives the item of her choice.

This design allows us to test whether subjects express demand for agency and whether it differs by gender. We also observe altruism in the baseline choice: since the items are gender-specific, altruistic husbands can select the purse and altruistic wives can select the wallet. Each spouse can also select the pencil case for one of their children. The novelty of this design is that it allows investigating demand for power, in terms of husbands' willingness to pay for power over their wife's consumption. We investigate this question by examining the proportion of husbands who refuse to delegate the choice of item to their wife, and by testing whether this refusal can be reversed by the offer of a monetary transfer.

The number of observations in each experimental cells is given in Table C.5 in Online Appendix C and power calculations are reported in Table C.6. Given the much larger sample size than in the two lab experiments, the minimum detectable effect size is smaller: 4 percentage points (pp) for treated-control comparisons, and 3.5 pp for testing the effect of the cost of agency. For delegation, the minimum detectable effect size is 0.9 pp. These smaller effect sizes are justified by the fact that, in a field experiment situation, even effects of such magnitude can be regarded as economically significant.

5.2 Empirical results

We first look at the gift selections made by husbands and wives. The purse is most popular, being selected by 81% of women and 64% of men. The wallet comes second, selected by 32% of men and 15% of women. Only 4% of men and women select the pencil case. There is considerable overlap between the choices made: 73% of couples select the same gift. Participants seem aware of this on average: 72% of women and 69% think their spouse has made the same selection as them. These

[37] Efforts were made during data collection to prevent communication between husband and wife while the two interviews were ongoing. Given the logistical difficulty of preventing communication between spouses in a field setting, we cannot however rule out that communication took place between some of the participants.

beliefs, however, are only true 59% to 63% of the time, respectively, and not aligned across spouses: in 43% of the couples at least one spouse believes that they have selected different things.[38] This creates space for instrumental value of agency, which we capture with a dummy M_i equal to 1 if individual i believes that her spouse has selected the same gift and 0 otherwise.

Only 16 husbands opt for the coin toss and refuse to fully delegate the choice of gift to their wife, providing little evidence in support of demand for power.[39]. Yet we find that a husband is significantly more likely to delegate when he believes his wife selected the same gift as him – a finding in line with the instrumental value of agency. It remains that 31% men believe their wife selected a different gift, and only a small fraction of them refuse to delegate the gift choice to their wife. In LAB1 and LAB2 we found that many subjects – including men – disregard the stated preferences of their spouse. Here we find overwhelming willingness to delegate. A possible interpretation is that process matters: if male subjects are explicitly asked to defer to their wife, many agree to do so – especially in the context of a female-oriented microfinance intervention. But if they are not asked to defer, they happily impose their own preferences.[40]

In Table 6 we regress a woman's decision to forfeit money in order to ensure she receives her preferred gift on the following regressors: M_i, which captures the instrumental value of agency; the cost of exerting agency; a dummy T_i if i's household was assigned to the microfinance treatment; and dummies for the preferred item. Since each woman makes this decision twice – once facing a price of PRK50 and once with PKR200 – the estimated regression combines both decisions.[41] Most women pay for agency: 77% when the price of agency is PKR50, 61% when it is PKR200 (i.e., 40% of the full cost of the gift).[42] We also find support for the instrumental value of agency: a woman who believes her husband chose the same gift as her is 4 percentage points less likely to pay for agency. However, the proportion of women paying for agency is much higher than the 28% of women who believe that their husband chose something else, indicating that much demand for agency is not instrumental.

Regarding the effect of being offered the microfinance treatment, we see that women invited to participate in the treatment are 4 percentage points less likely to demand agency.[43] In other words, demand

[38] Specifically: in 57% of couples both spouses believe their selections are identical; in 17% both believe they each selected different things; and in 27% spouses have misaligned beliefs, i.e., one spouse thinks the other has selected the same thing, and the other thinks the opposite.

[39] This is below the minimum detectable size of 0.9 percentage points reported in Table C.6

[40] Appendix A Table A6 shows regression results of the husband's refusal to delegate on M_i, and a dummy T_i if i's household was assigned to the microfinance treatment and 0 otherwise.

[41] To allow for possible correlation between the two agency decisions, standard errors are clustered at the individual level.

[42] This difference of 16 percentage points is much larger than the minimum detectable effect of 3.5 percentage points reported in Table C.6.

[43] This is just within the minimum detectable effect size calculated in Table C.6. Heterogeneity analysis shows that this effect is driven by women with low decision-making power at baseline.

for agency is high among participating women, but women empowered by the microfinance intervention are slightly less inclined to pay for agency. If we combine this result with the higher demand for power and agency among women documented in LAB1 and LAB2, it is consistent with the following interpretation: the women in our setting – whose decision-making autonomy is severely limited in their daily lives – express pent-up demand for agency within the safety of our experiments; but access to finance, by empowering them, reduces that pent-up demand slightly (Ashraf et al., 2010). We do, however, acknowledge that these findings are not sufficient, by themselves, to *prove* a reduced pent-up demand interpretation. Other factors may play a role as well – e.g., access to finance may change the household's budget constraint over and above its effect on female empowerment and women's outside options.[44] Moreover, we find no evidence that women with greater decision-making autonomy within the household are less likely to exert power and agency during our laboratory experiments.

6 Conclusion

In this paper we have tested specific questions regarding demand for agency and power within and across households using two lab experiments and one field experiment. We have investigated whether individuals know and respect the preferences of others, and whether subjects are willing to pay for control over their own consumption.

The evidence we uncover suggests that the relationship between gender and agency is complex. First, the fact that household members often ignore – or choose to ignore – each other's consumption preferences is a surprising finding which, a priori, suggests that individual executive agency over own consumption has an instrumental value. While demand for agency responds to the benefits and costs of agency among both men and women, we find female subjects in our study population to be equally or more willing to exert agency than their male counterparts. This is true in spite of the fact that, when asked about executive and consultative agency within the household, our study population appears strongly patriarchal.

What have we learned about demand for agency? Men and women seem to value making consumption choices for themselves and others, over and beyond what these choices imply for their own material consumption. The process by which consumption is decided seems to be an important consideration for intrahousehold welfare. To an economist, these tendencies are disturbing because they imply that, even in relatively poor households such as those we study, people are willing to reduce allocative

[44] This is similar to the issue raised by Attanasio and Lechene (2014) when discussing the impact of PROGRESA. We thank an anonymous reviewer for pointing this out. We should however point out that Afzal et al. (2020) find no effect of the RCT on income, consumption, or assets, thereby mitigating this concern in our case.

efficiency in consumption simply to assuage a desire for power and agency. More research is needed on preferences over process in intrahousehold decisions and on their effect on material and subjective welfare.

References

Afzal, U., G. d'Adda, M. Fafchamps, S. Quinn, and F. Said (2018). Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan. *The Economic Journal 128*(614), 2161–2190. 5.1

Afzal, U., G. d'Adda, M. Fafchamps, S. Quinn, and F. Said (2019). When Nudge Turns to Shove: A Field Experiment Adding Commitment Features to a Commitment Savings Product. mimeo, Oxford University. 1, 5.1

Alan, S., S. Ertac, and I. Mumcu (2018). Gender Stereotypes in the Classroom and Effects on Achievement. *The Review of Economics and Statistics 100*(5), 876–890. 2

Ambler, K. (2015). Don't tell on me: Experimental evidence of asymmetric information in transnational households. *Journal of Development Economics 113*, 52–69. 1

Anderson, S. and J. Baland (2002). The Economics of ROSCAs and Intrahousehold Resource Allocation. *The Quarterly Journal of Economics 117*(3), 963–995. 1

Andrews, I. and M. Kasy (2019, August). Identification of and correction for publication bias. *American Economic Review 109*(8), 2766–94. 1

Angelucci, M. and R. Garlick (2016). Heterogeneity in the Efficiency of Intrahousehold Resource Allocation: Empirical Evidence and Implications for Investment in Children. Working paper. 1

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the Philippines. *American Economic Review 99*(4), 1245–1277. 1

Ashraf, N., E. Field, and J. Lee (2014). Household Bargaining and Excess Fertility: An Experimental Study in Zambia. *American Economic Review 104*, 2210–2237. 1

Ashraf, N., D. Karlan, and W. Yin (2010). Female Empowerment: Impact of a Commitment Savings Product in the Philippines. *World Development 38*(3), 333–344. 5.2

Bakhtiar, M. (2019). Sharing the Pie: Undernutrition, Intra-household Allocation, and Poverty. Unpublished phd thesis, University of Maryland. 1

Baland, J.-M. and R. Ziparo (2017). Intra-household bargaining in poor countries. WIDER Working Paper Series wp-2017-108, World Institute for Development Economic Research (UNU-WIDER). 1

Banerjee, A., E. La Ferrara, and V. Orozco (2018). The Entertaining Way to Behavioral Change: Fighting HIV with MTV. mimeo. 2

Barr, A., M. Dekker, W. Janssens, B. Kebede, and B. Kramer (2019). Cooperation in Polygynous Households. *American Economic Journal: Applied Economics 11*(2), 266–83. 1

Bartling, B., E. Fehr, and H. Herz (2014). The Intrinsic Value of Decision Rights. *Econometrica 82*(6), 2005–2039. 1

Becker, G. S. (1981). *A Treatise on the Family*. Cambridge Mass.: Harvard UP. 2

Behrman, J. (1993). Intrahousehold distribution and the family. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1, Part A, Chapter 04, pp. 125–187. Elsevier. 1

Bernard, T., S. Dercon, K. Orkin, and A. Taffesse (2016). The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia. CSAE Working Paper Series 2014-16, Centre for the Study of African Economies, University of Oxford. 2

Bloch, F. and V. Rao (2002). Terror as a bargaining instrument: A case study of dowry violence in rural India. *American Economic Review 92*(4), 1029–1043. 2

Bobonis, G. (2009). Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment. *Journal of Political Economy 117*(3), 453–503. 3

Bolton, G., J. Brandts, and A. Ockenfels (2005). Fair Procedure: Evidence from Games Involving Lotteries. *The Economic Journal 115*, 1054–1076. 1

Brown, C., R. Calvi, and J. Penglase (2020). Sharing the Pie: Undernutrition, Intra-household Allocation, and Poverty. mimeo. 1

Browning, M., F. Bourguignon, P.-A. Chiappori, and V. Lechene (1994). Income and Outcomes: A Structural Model of Intrahousehold Allocation. *Journal of Political Economy 102*(6), 1067–1096. 1

Browning, M. and P.-A. Chiappori (1998). Efficient Intra-Household Allocations: A General Characterization and Empirical Tests. *Econometrica 66*(6), 1241–1278. 1, 2

Browning, M., P.-A. Chiappori, and V. Lechene (2010). Efficient Intra-Household Allocations: A General Characterization and Empirical Tests. *Economic Journal 120*(545), 786–799. 1

Browning, M., P.-A. Chiappori, and A. Lewbel (2013). Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power. *Review of Economic Studies 80*(4), 1267–1303. 1

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge: Cambridge University Press. 1, 2

Camerer, C. F., A. Dreber, E. Forsell, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, and T. Chan (2016). Evaluating replicability of laboratory experiments in economics. *Science 351*(6280), 1433–36. 1

Carlsson, F., P. Martinsson, P. Qin, and M. Sutter (2009). Household Decision Making and the Influence of Spouses' Income, Education, and Communist Party Membership: A Field Experiment in Rural China. Scholarly Paper 1395246, SSRN. 1

Castilla, C. (2014). What's yours is mine, and what's mine is mine: Bargaining power and income concealing between spouses in India. Working paper. 1

Castilla, C. (2015). Trust and Reciprocity between Spouses in India. *American Economic Review 105*(5), 621–624. 1

Castilla, C. and T. Walker (2013). Trust and Reciprocity between Spouses in India. *American Economic Review 103*, 263–268. 1

Chen, J. (2013). Identifying Non-cooperative Behavior among Spouses: Child Outcomes in Migrant-sending Households. *Journal of Development Economics 100*(1), 1–18. 1

Chiappori, P.-A. (1997a). Intrahousehold resource allocation in developing countries: Methods, models and policy. In J. H. Lawrence Haddad and H. Alderman (Eds.), *Collective Models of Household Behavior: The Sharing Rule Approach*. IFPRI, Washington D.C. 1

Chiappori, P.-A. (1997b). Introducing Household Production in Collective Models of Labor Supply. *Journal of Political Economy 105*(1), 191–209. 1

Christensen, G. and E. Miguel (2016). Transparency, Reproducibility, and the Credibility of Economics Research. NBER Working Papers 22989, National Bureau of Economic Research, Inc. 1

de Brauw, A., D. Gilligan, J. Hoddinott, and S. Roy (2014). The Impact of Bolsa Família on Women's Decision-Making Power. *World Development 59*, 487–504. 3

de Laat, J. (2014). Household Allocations and Endogenous Information: The case of split migrants in Kenya. *Journal of Development Economics 106*, 108–117. 1

Dean, J. and S. Jayachandran (2019). A Family Affair: Family Members' Role in Female Employment Decisions in India. mimeo., Northwestern University. 2

Deaton, A. and C. Paxson (1998, October). Economies of Scale, Household Size, and the Demand for Food. *Journal of Political Economy 106*(5), 897–930. 1

Deer, C. D. and J. Twyman (2012). Asset Ownership and Egalitarian Decision Making in Dual-headed Households in Ecuador. *Review of Radical Political Economics 44*(3), 313–320. 1

Dercon, S. and P. Krishnan (2000). In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia. *Journal of Political Economy 108*(4), 688–727. 1

Duflo, E. and C. Udry (2004, May). Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices. Working Paper 10498, National Bureau of Economic Research. 1

Eckel, C. (2016). Making Decisions for Others: The Consequences of Gender Stereotypes. mimeo., Texas A & M. 4.2

Fafchamps, M., B. Kebede, and A. R. Quisumbing (2009). Chivalry and Solidarity in Ultimatum Games. *Oxford Bulletin of Economics and Statistics 71*(4), 567–599. 1

Falk, A. and M. Kosfeld (2006). The Hidden Cost of Control. *The American Economic Review 96*(5), 1611–1630. 4

Fehr, E., H. Herz, and T. Wilkening (2013). The Lure of Authority: Motivation and Incentive Effects of Power. *The American Economic Review 103*(4), 1325–1359. 1

Ferro-Luzzi, A., C. Scaccini, S. Taffese, B. Aberra, and T. Demeke (1990). Seasonal Energy Deficiency in Ethiopian Rural Women. *European Journal of Clinical Nutrition 44*(1), 7–18. 1

Fiala, N. (2017). Business is tough, but family is worse: The role of family constraints on microenterprise development in Uganda. mimeo., The World Bank. 1

Field, E. and K. Vyborny (2016). Female Labor Force Participation: Pakistan Country Study for the Asian Development Bank. Policy brief, Asian Development Bank. 4.1

Güth, W. and H. Weck-Hannemann (1997). Do People Care about Democracy? An Experiment Exploring the Value of Voting Rights. *Public Choice 91*, 27–47. 1

Haddad, L., J. Hoddinott, and H. Alderman (1997). *La Pensee Sauvage*. Baltimore and London: Johns Hopkins University Press. 1

Hoddinott, J. and L. Haddad (1995, February). Does Female Income Share Influence Household Expenditures? Evidence from Cote d'Ivoire. *Oxford Bulletin of Economics and Statistics 57*(1), 77–96. 1

Hoel, J. B. (2015). Heterogeneous Households: A Within-subject Test of Asymmetric Information between Spouses in Kenya. *Journal of Economic Behavior & Organization 118*, 123–135. 1

Iversen, V., C. Jackson, B. Kebede, A. Munro, and A. Verschoor (2006). What's Love Got To Do With It? An Experimental Test of Household Models in East Uganda. Discussion Papers in Economics 06/01, Royal Holloway University of London. 1

Iversen, V., C. Jackson, B. Kebede, A. Munro, and A. Verschoor (2011). Do Spouses Realise Cooperative Gains? Experimental Evidence from Rural Uganda. *World Development 39*(4), 569–578. 1

Jakiela, P. and O. Ozier (2015). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. *Review of Economic Studies 83*(1), 231–268. 1

Kebede, B., M. Tarazona, A. Munro, and A. Verschoor (2014). Intra-household Efficiency: An Experimental Study from Ethiopia. *Journal of African Economies 23*(1), 105–150. 1

Krishnan, P. and S. Krutikova (2013). Non-cognitive skill formation in poor neighbourhoods of urban India. *Labour Economics 24*, 68–85. 2

Lundberg, S. and R. A. Pollak (1996). Bargaining and Distribution in Marriage. *Journal of Economic Perspectives 10*(4), 139–158. 1

Mani, A. (2011). Mine, yours or ours? The efficiency of household investment decisions: An experimental approach. Working paper. 1

McElroy, M. B. (1990). The Empirical Content of Nash-Bargained Household Behavior. *Journal of Human Resources 25*(4), 559–583. 1

Munro, A., I. Bateman, and T. McNally (2008). The Family Under the Microscope: An Experiment Testing Economic Models of Household Choice. Scholarly Paper 1138969, SSRN. 3

Owens, D., Z. Grossman, and R. Fackler (2014). The Control Premium: A Preference for Payoff Autonomy. *American Economic Journal: Microeconomics 91*, 138–161. 1

Pitt, M. M., M. R. Rosenzweig, and M. Hassan (1990, December). Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. *American Economic Review 80*(5), 1139–1156. 1

Pollak, R. A. (2004). An Intergenerational Model of Domestic Violence. *Journal of Population Economics 17*(2), 311–329. 2

Rangel, M. and D. Thomas (2019). Decision-Making in Complex Households. NBER Working Papers 26511, National Bureau of Economic Research, Inc. 1

Rao, V. (1997). Wife-beating in Rural South India: A Qualitative and Econometric Analysis. *Social Science & Medicine 44*(8), 1169–1180. 2

Robinson, J. (2012). Limited Insurance within the Household: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics 4*(4), 140–164. 3

Schaner, S. (2012). Do Opposites Detract? Intrahousehold Preference Heterogeneity and Inefficient Strategic Savings. *American Economic Journal: Applied Economics 7*(2), 135–174. 1

Strauss, J. and D. Thomas (1998). Health, nutrition, and economic development. *Journal of Economic Literature 36*(2), 766–817. 1

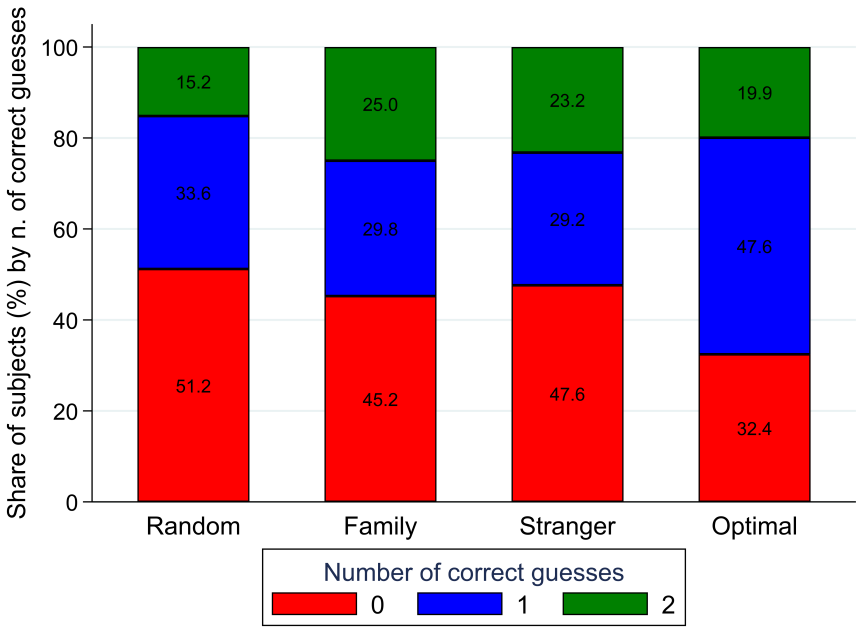
Thomas, D. (1994). Like father, like son; like mother, like daughter: Parental resources and child height. *Journal of Human Resources 29*(4). 1

Verschoor, A., B. Kebede, A. Munro, and M. Tarazona (2019). Spousal Control and Efficiency of Intrahousehold Decision Making: Experiments among Married couples in India, Ethiopia and Nigeria. *European Journal of Development Research forthcoming*. 1

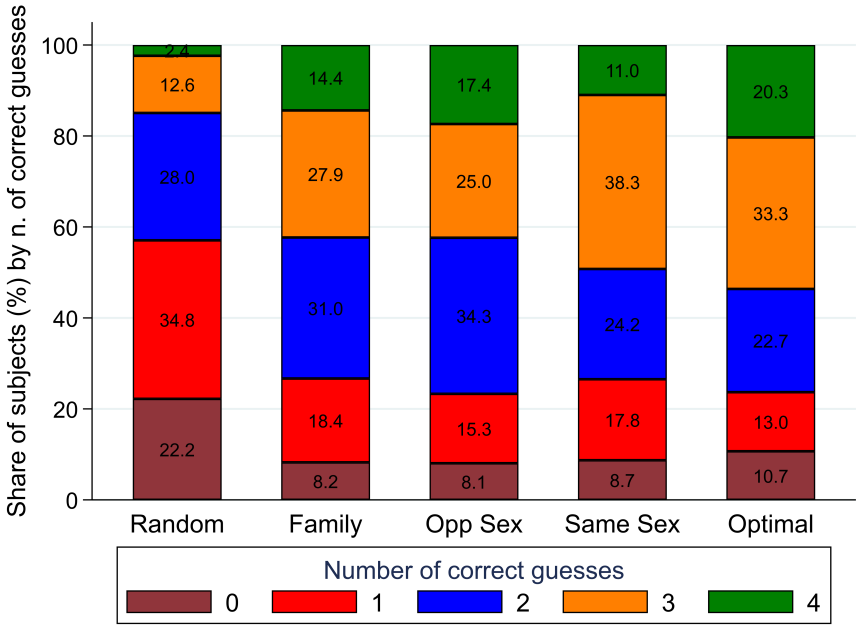
World Economic Forum (2020). Global Gender Gap Report 2020. Technical report, World Economic Forum. 4.1

Figures

Figure 1: Guess quality by match type in LAB1 and LAB2



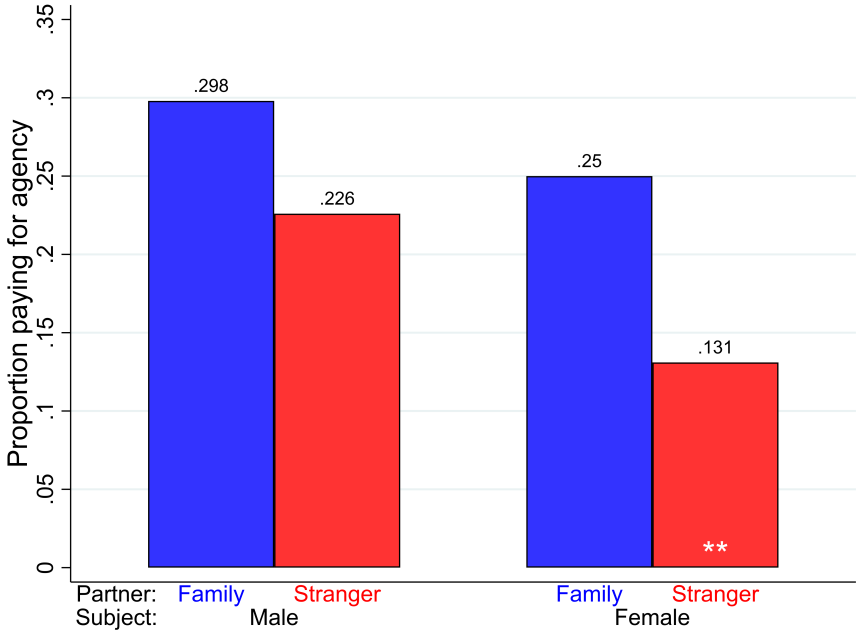
Panel (a): LAB1



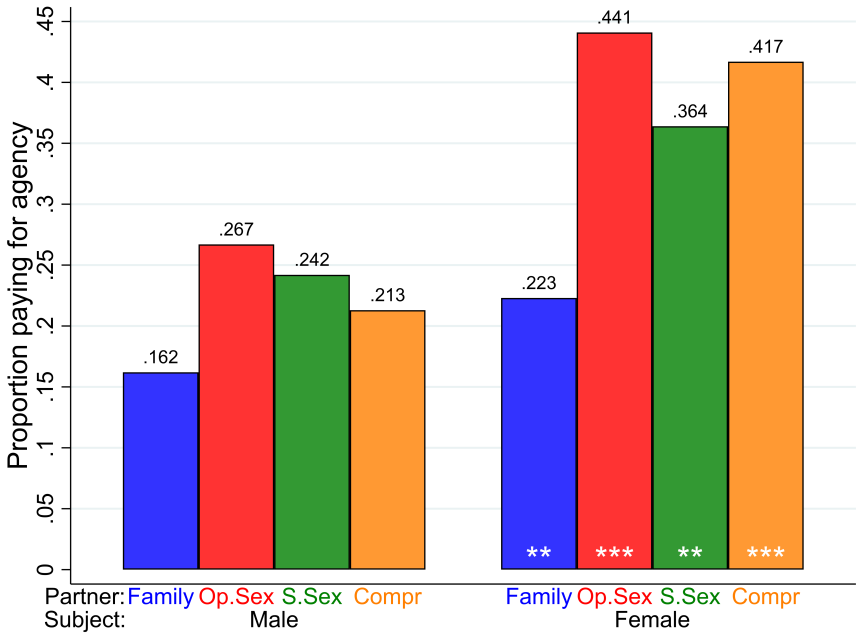
Panel (b): LAB2

Notes: Figure 1 shows the distribution of guess quality for different types of guesses. Panel (a) displays guess quality for LAB1, where guess quality takes three values: 2 for correctly guessing the top and bottom ranked juice; 1 for correctly guessing either the top or bottom ranked juice; and 0 for guessing neither. In Panel (b) guess quality takes five values: 4 for correctly guessing the top and bottom ranked food and drink; 1, 2 or 3 for making 1, 2 or 3 of these four guesses correctly; and 0 for not getting any of the four guesses right. In each panel, the bars depict the distribution of guess quality of lab participants, depending on their matching treatment, with bar values specified in the middle of the bar. The Random bar depicts the hypothetical distribution of guess quality of a subject who guesses at random. The Optimal bar depicts the hypothetical distribution of guess quality of an informed subject who guesses optimally, that is, who picks the mode of the sample distribution as their best guess.

Figure 2: Demand for ex ante agency in LAB1 and LAB2



Panel (a): LAB1



Panel (b): LAB2

Notes: The bars report the frequency with which participants of both genders pay for agency when matched with different partners, with proportions presented on top of the bar. In Panel (a), we display the frequency for LAB1. The first bar considers men matched with their wife or female family member; the second, men matched with a stranger of the opposite gender; the third and the fourth bars likewise present proportions for women by match type. In Panel (b), we display the frequency in LAB2. The first bar corresponds to the frequency with which male participants pay for agency when matched with their wife or female family member. The other bars correspond to the other combination of gender and matching treatment. In each panel, stars report the significance level of a test of equality with the bar of the same color, e.g., is Female-Stranger equal to Male-Stranger.

Tables

Table 1: Descriptive statistics on the RCT and LAB samples

	LAB1		**LAB2**		**FIELD**	
	n	*mean*	*n*	*mean*	*n*	*mean*
Age	336	38.3	716	36.1	1991	39.4
Can read and write	324	0.63	716	0.48	1991	0.49
Average monthly household expenses (PKR)	336	14,406	716	14,291	1991	18,474
Self employed (females)	156	0.26	358	0.14	1991	0.11
Housewives (females)	156	0.37	358	0.87	1991	0.74
Head of the household (females)	168	0.10	358	0.00	1991	0.12
Spouse of the household head (females)	168	0.74	358	0.73	1991	0.74
N. decisions where permission is required (males)	168	1.13	358	4.89	–	–
N. decisions where permission is required (females)	168	5.57	358	8.03	1991	5.13
N. decisions opinion taken into account (females)	–	–	–	–	1991	9.78

Note: For LAB1 and LAB2 we report the average age and literacy for participating subjects of both sexes, as well as their joint household income. For the field experiment, we report the average age and literacy of the female respondent, as well as her household income.'N. decisions where permission required' refers to number of decisions, out of 9, where the respondent requires permission from spouse. In field experiment, this question is only administered to females. In the field experiment, we also ask about the number of decisions, out of a total of 14, where female opinions are taken into account. Gender of the respondent is specified in parentheses.

Table 2: Respect for partner's preferences

Dependent variable	Probability of choosing ...							
	...one's own guess of the partner's preferred option				...the partner's revealed preferred option			
Treatment sub-sample	No information				Information			
Experiment	LAB1		LAB2		LAB1		LAB2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male – Stranger	-0.125	-0.117	-0.033	-0.031	-0.253**	-0.262**	-0.079	-0.102**
Opposite Sex	(0.102)	(0.103)	(0.043)	(0.044)	(0.108)	(0.109)	(0.052)	(0.052)
Male – Stranger			-0.066	-0.061			-0.038	-0.060
Same Sex			(0.047)	(0.050)			(0.050)	(0.050)
Female – Family	-0.186**	-0.173**	-0.098***	-0.101***	-0.049	-0.059	-0.123***	-0.124***
	(0.077)	(0.080)	(0.030)	(0.030)	(0.078)	(0.078)	(0.031)	(0.030)
Female – Stranger	-0.173*	-0.162	-0.0667	-0.0664	-0.0621	-0.0606	-0.0620	-0.0614
Opposite Sex	(0.103)	(0.104)	(0.044)	(0.044)	(0.110)	(0.110)	(0.051)	(0.051)
Female – Stranger			-0.0282	-0.0294			-0.0684	-0.0796
Same Sex			(0.045)	(0.047)			(0.052)	(0.053)
Own altruism		0.077		-0.021		-0.050		0.055
towards partner		(0.159)		(0.048)		(0.171)		(0.053)
(Expected) partner's		-0.027		0.041		0.151		-0.144**
altruism		(0.148)		(0.048)		(0.165)		(0.060)
Item is a drink			0.036	0.036			0.049*	0.049*
			(0.022)	(0.022)			(0.025)	(0.025)
Intercept	0.744***	0.710***	0.828***	0.819***	0.610***	0.560***	0.741***	0.787***
(Male – Family)	(0.068)	(0.147)	(0.023)	(0.045)	(0.077)	(0.161)	(0.029)	(0.048)
N. Observations	170	170	1216	1216	166	166	1216	1216

Note: Standard errors clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%. LAB2 observations include family matching observations plus stranger matching; subjects matched with a computer are not included since computers have no preferences. Furthermore, observations for food and drinks are combined. This means in LAB2 there are four observations per subject: food and drink for family matching; and food and drink for stranger matching – except for subjects matched with a computer, in which case there are two observations per subjects: food and drink for family matching. For uninformed subjects, the dependent variable equals 1 if the subject picks their best guess of their partner's top ranked item, and 0 otherwise. For informed subjects, the dependent variable equals 1 if the subject picks what they know to be their partner's top ranked item, and 0 otherwise. In LAB1, the number of informed and uninformed subjects differs slightly due to variation in the number of subject pairs across sessions.

Table 3: Testing demand for ex ante instrumental agency

Dependent variable	Ex-ante agency: paying to get own pick before knowing partner's pick for self			
Experiment	LAB1		LAB2	
	(1)	(2)	(3)	(4)
Expected gain from choosing own pick: Drink		-0.069		-0.045
		(0.261)		(0.048)
Expected gain from choosing own pick: Food				0.136**
				(0.058)
Cost of getting own pick			-0.001	-0.001*
			(0.001)	(0.001)
Male – Stranger Opposite Sex	-0.048	-0.044	0.099**	0.094**
	(0.069)	(0.068)	(0.045)	(0.045)
Male – Stranger Same Sex			0.080*	0.087**
			(0.042)	(0.042)
Male – Computer			0.053	0.034
			(0.044)	(0.047)
Female – Family	-0.071	-0.068	0.063**	0.070**
	(0.070)	(0.070)	(0.028)	(0.029)
Female – Stranger Opposite Sex	-0.167***	-0.160**	0.281***	0.285***
	(0.063)	(0.065)	(0.050)	(0.051)
Female – Stranger Same Sex			0.203***	0.207***
			(0.046)	(0.046)
Female – Computer			0.256***	0.241***
			(0.052)	(0.053)
Intercept (Male – Family)	0.298***	0.312***	0.184***	0.153***
	(0.050)	(0.078)	(0.025)	(0.031)
N. Observations	336	336	1432	1432

Note: Standard errors are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports IV estimates of the effect of the expected gain from agency on willingness to pay for agency. In LAB1 willingness to pay for agency equals 1 if the subject prefers half a glass of their preferred to a coin toss. In LAB2 willingness to pay for agency equals 1 if the subject forfeits a fraction of their show-up fee to receive their top food and drink instead of receiving the food and drink selected by their partner. At the time of making the decision, subjects do not know what the partner has selected for them. Instruments include a fully interacted set of dummies for gender, informed treatment, matching type, and own top rank. In LAB1 the predicted gain in the probability of getting preferred juice varies between 8% and 44%, with a mean of 25% (s.d. 9%). In LAB2 the predicted gain vary between 8% and 100% for both food and drink, with an average of 36% (s.d. 24%) for food and 34% (s.d. 26%) for drink. The cost of agency takes values 5, 20 and 50, equally randomized across subjects. For all matching types, demand for agency in LAB2 tests as significantly higher for female than male subjects.

Table 4: Proportion of LAB2 subjects who demand ex post agency

Dependent variable	Ex post agency: proportion of subjects who demand agency after being informed of partner's pick				
Sub-sample	Wrong food - Wrong drink (1)	Wrong food - Right drink (2)	Right food - Wrong drink (3)	Right food - Right drink (4)	All (5)
Panel A: overall					
All	17.2% (250)	20.5% (269)	12.2% (230)	10.8% (683)	14.0% (1432)
Female	16.7% (114)	18.4% (125)	11.5% (104)	8.3% (373)	11.9% (716)
Male	17.7% (136)	22.2% (144)	12.7% (126)	13.9% (310)	16.1% (716)
Panel B: conditional on ex-ante agency					
Exerted agency ex ante	52.9% (68)	61.8% (76)	40.7% (59)	35.5% (166)	45.0% (369)
Did not exert agency ex-ante	3.9% (182)	4.2% (193)	2.3% (173)	2.9% (517)	3.2% (1063)

Note: Each percentage is the proportion of LAB2 subjects willing to pay for agency after having been informed of their partner's food and drink selection for them. Number of respondents in each subgroup is reported in parenthesis. Subjects only make one decision combining food and drink. The top panel reports these proportions for all subjects and by gender, depending on whether the partner selected the correct food and drink or not. The bottom panel breaks down these proportions between subjects who demanded agency ex ante – i.e., before being informed of their partner's selection for them – and those who did not.

Table 5: Testing demand for ex post pure agency

Dependent variable	Ex-post agency: paying to get own pick after knowing partner's pick for self
Experiment	LAB2
	(1)
Actual gain from choosing own pick: Drink	-0.004
	(0.021)
Actual gain from choosing own pick: Food	0.077***
	(0.021)
Cost of getting own pick	-0.002***
	(0.001)
Male – Stranger Opposite Sex	0.138***
	(0.044)
Male – Stranger Same Sex	0.066*
	(0.037)
Male – Computer	0.025
	(0.039)
Female – Stranger Opposite Sex	-0.017
	(0.022)
Female – Stranger Same Sex	0.039
	(0.038)
Female – Stranger Same Sex	-0.008
	(0.032)
Female – Computer	0.050
	(0.040)
Intercept (Male – Family)	0.140***
	(0.022)
N. Observations	1432

Note: Standard errors are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports OLS estimates of the effect of the actual gain from agency on willingness to pay for agency. The dependent variables equals 1 if the subject forfeits a fraction of the showup fee to receive their top ranked food and drink, and 0 otherwise. At the time of making the decision, subjects know what the partner has selected for them. The cost of agency takes values 5, 20 and 50, equally randomized across subjects.

Table 6: Testing demand for ex ante instrumental agency in the field experiment

Dependent variable	Ex ante agency: willingness to pay to get own choice of gift for sure
Experiment	Field experiment
	(1)
Low expected gain from agency (own pick - expected husband's pick)	-0.039**
	(0.020)
High cost of agency	-0.160***
	(0.011)
Treated in the RCT	-0.039**
	(0.020)
Own pick = Wallet	0.001
	(0.025)
Own pick = Pencil case	0.178***
	(0.034)
Intercept	0.814***
	(0.023)
N. Observations	3982

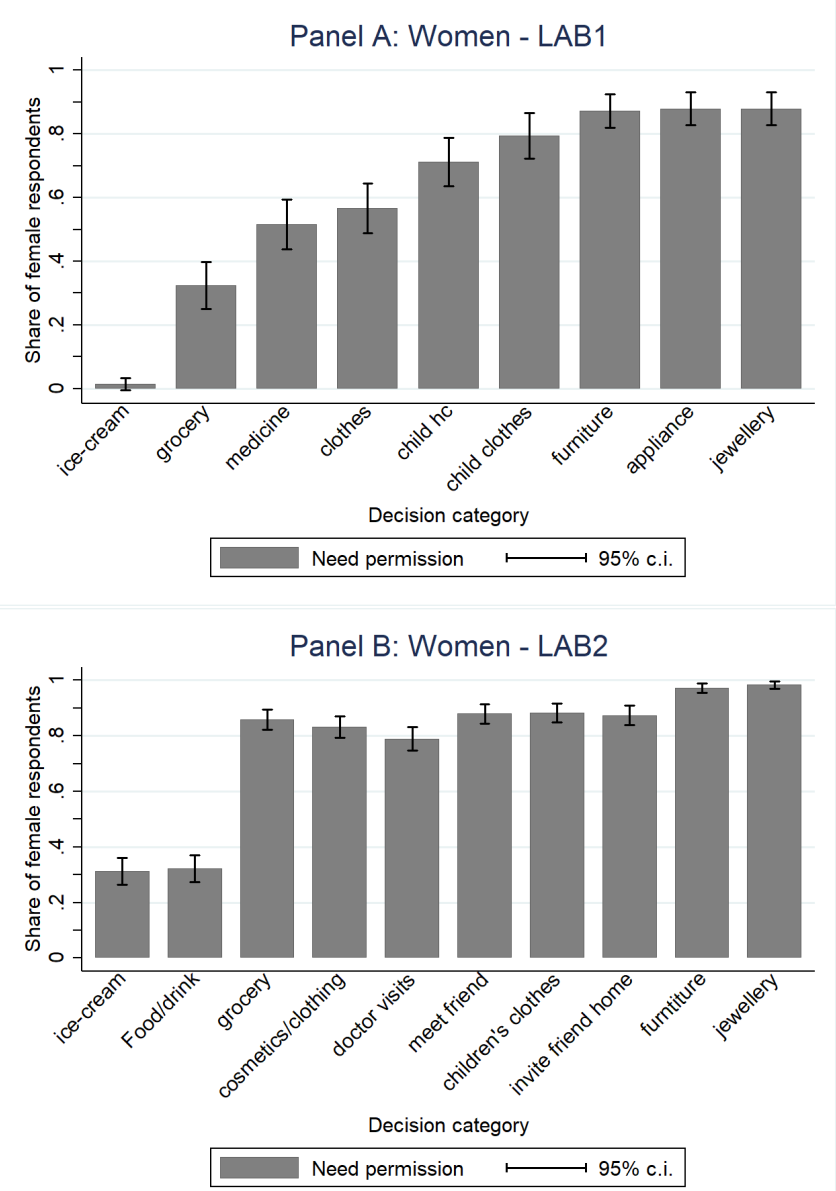
Note: Standard errors are clustered at the subject level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports OLS estimates of the value and cost of agency on willingness to exert agency. The dependent variables equals 1 if the subject is willing to forfeit part of a monetary participation fee, and 0 otherwise. Subjects make two separate decisions, one for which the cost of agency is PKR50 (the default); and the other for PKR200 (High cost of agency dummy=1). At the time of making the decision, subjects do not know what the partner has selected for them, but they have been asked to make a guess. The main regressor of interest equals 1 if the preference of the wife is equal to the choice she expects her husband to make.

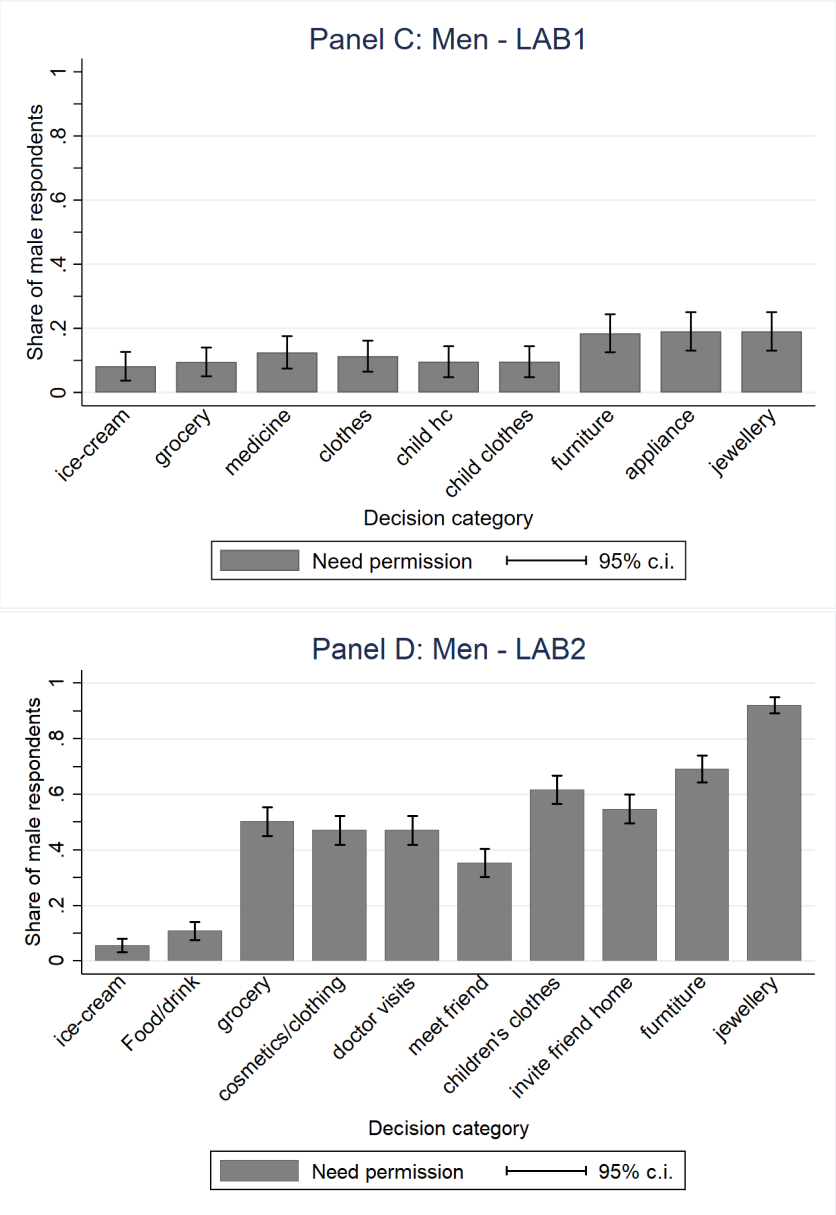
The sample consists of 1991 male and 1991 female respondents.

Appendix A: Additional tables and figures

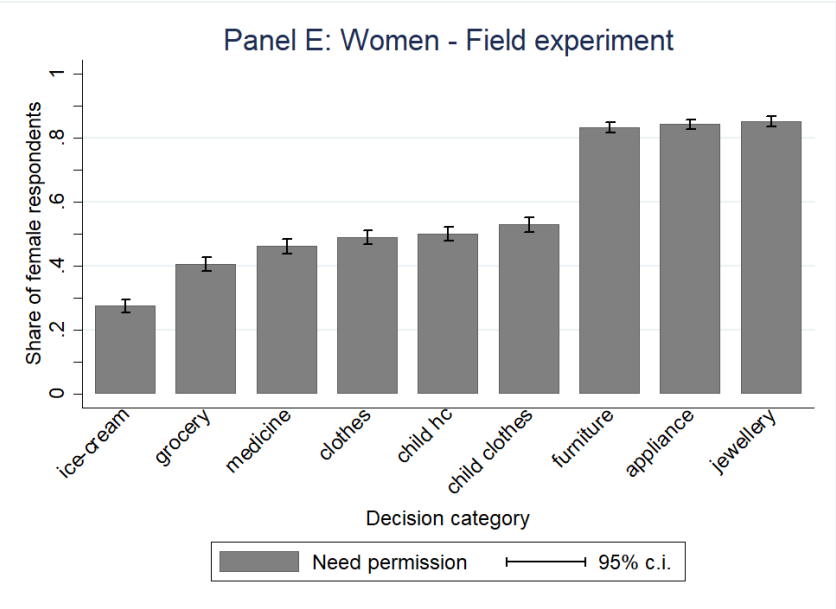
Figure A1: Executive agency:
Share of respondents who need permission from spouse



Notes: The bar labelled 'jewellery' represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *sell* jewellery?'. In LAB2 the bar labelled 'invite friend home' presents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *invite a friend home*?' All the other bars represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *buy* the listed item?' Whiskers show the 95% confidence interval for each bar.

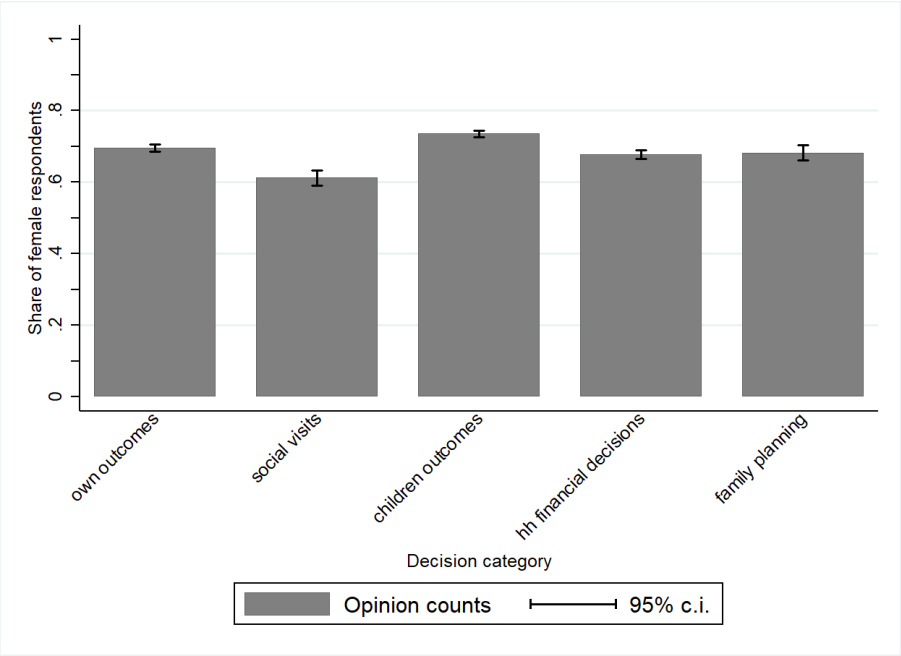


Notes: The bar labelled 'jewellery' represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *sell* jewellery?'. In LAB2 the bar labelled 'invite friend home' presents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *invite a friend home*?' All the other bars represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *buy* the listed item?' Whiskers show the 95% confidence interval for each bar.



Notes: The bar labelled 'jewellery' represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *sell* jewellery?'. In LAB2 the bar labelled 'invite friend home' presents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *invite a friend home*?' All the other bars represents the percentage of respondents who answer 'yes' to the question 'Do you need permission from your spouse to *buy* the listed item?' Whiskers show the 95% confidence interval for each bar.

Figure A2: Consultative agency:
Share of women whose opinion is taken into account by husband



Notes: Each bar represents the percentage of female respondents in the field experiment who answer 'yes' to the question 'Does your husband take your opinion into account in decisions regarding [a particular decision]? The 'own outcomes' decision category is the average of answers to questions on decisions regarding the respondent's: medical care; ability to work for earned income, and ability to borrow money from an MFI. The 'social visits' decision category is the average of answers to questions on decisions to visit various types of family members. The 'children outcomes' category is the average of answers to questions on: boys' schooling; girls' schooling; children medical care; and the marriage of the children. The 'financial decisions' category is the average of answers to questions on: the purchase of electronic appliances for the household; decisions about house repairs; and decisions regarding the sale or purchase of a house. The 'family planning' category covers decisions about family planning. Whisker bars show the 95 confidence interval for each bar.

Table A1: Ability to guess partners' preferences by matching type

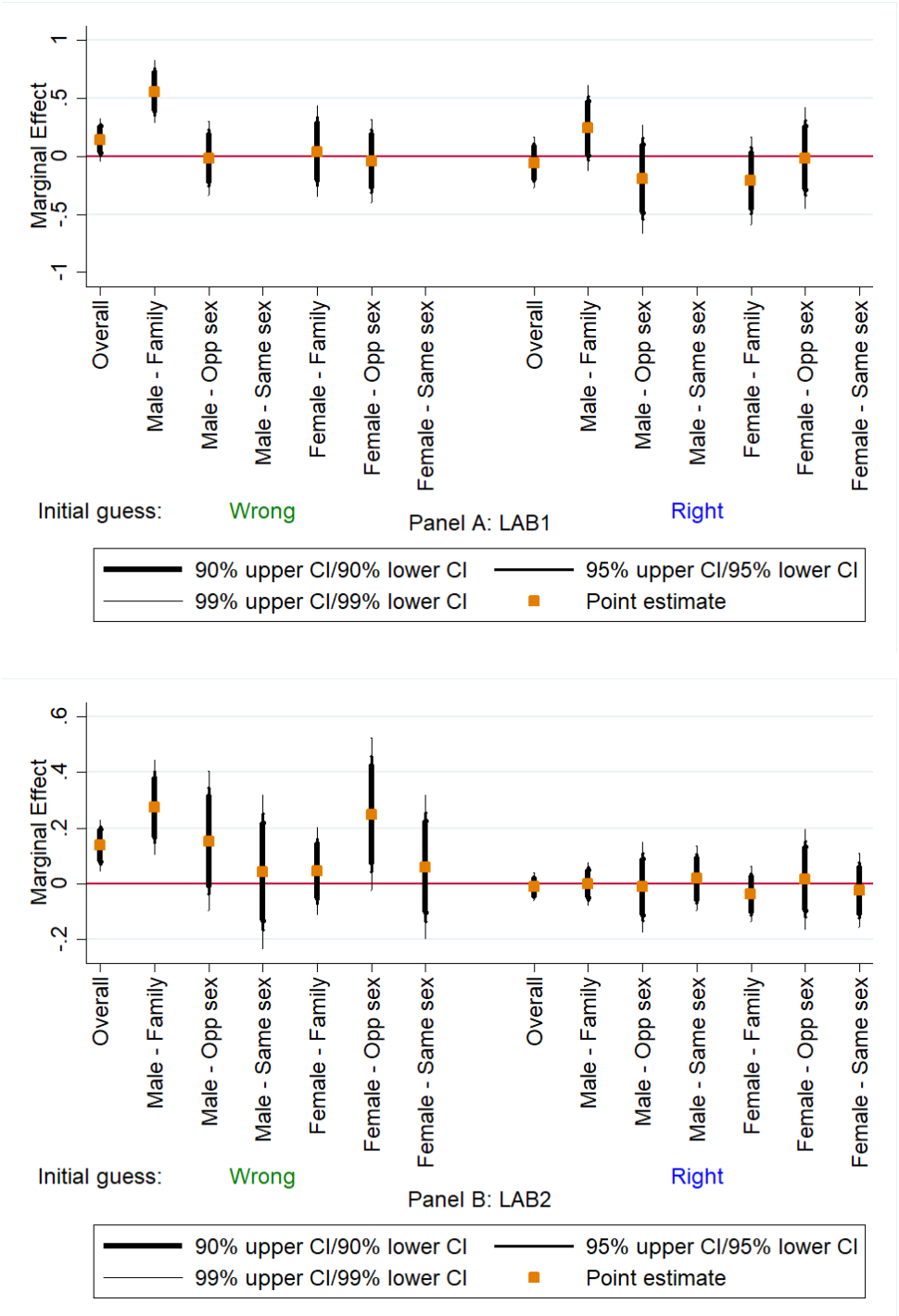
Dependent variable	Quality of guess	
Experiment	LAB1	LAB2
	(1)	(2)
Match type:		
Family member	0.158**	0.837***
	(0.071)	(0.052)
Stranger of opposite sex	0.116	0.903***
	(0.071)	(0.079)
Stranger of same sex		-0.436***
		(0.082)
Computer		1.015***
		(0.042)
Optimal uninformed benchmark	0.235***	0.869***
	(0.058)	(0.076)
Female	0.095**	-0.043
	(0.047)	(0.034)
Intercept	0.592***	1.402***
	(0.047)	(0.034)
Observations	1008 (336)	4296 (1432)

Standard errors are clustered at the pair level and in parenthesis. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

Notes: The dependent variable is the quality of the guess. In LAB1 it equals 1 if i correctly guesses j's top and bottom ranked juice. In LAB2 it equals 4 if i has 4 correct guesses (i.e., correctly guesses j's top ranked food and drink and bottom ranked food and drink), 3 if i has 3 correct guesses, etc. The omitted match category is an artificial set of observations in which each subject guesses the partner's preferences using population frequencies as probabilities. As a result, coefficients can be interpreted as improvement over random guessing, e.g., the intercept is the average quality of a random guess, and the coefficient of spouse/family member shows by the improvement of the guess relative to random guessing. We also include optimal guessing by an uninformed subject as benchmark to facilitate comparisons. Note that this regression corresponds to one specific vector of random guesses. The absolute value of coefficients changes with each draw, but relative values do not. Possible variation in the significance of coefficients (relative to random) is discussed in the text.

The last row shows the number of observations and the number of subjects in parenthesis. In LAB1 there are 336 subjects. For each of these observations, one random guess and one optimal uninformed guess are then added, thereby adding 336 x 2 fictitious observations. In LAB2 subjects appear twice: once matched with their family member, and once with a stranger, thereby making 1432 observations. For each of these observations, one random guess and one optimal uninformed guess are then added, thereby adding 1432 x 2 fictitious observations. The role of these fictitious observations is to allow an easy comparison of actual guesses to these two benchmarks.

Figure A3: Marginal effect of information on choice, conditional on guessing the partner's top rank



Notes: Marginal effects are estimated using a regression of the dependent variable on a fully interacted set of an informed treatment dummy, a gender dummy, and dummies for each matching type. The regression is estimated separately for subjects who incorrectly guessed the top rank of their partner, and subjects who guessed correctly. Regression results are shown in Table A2.

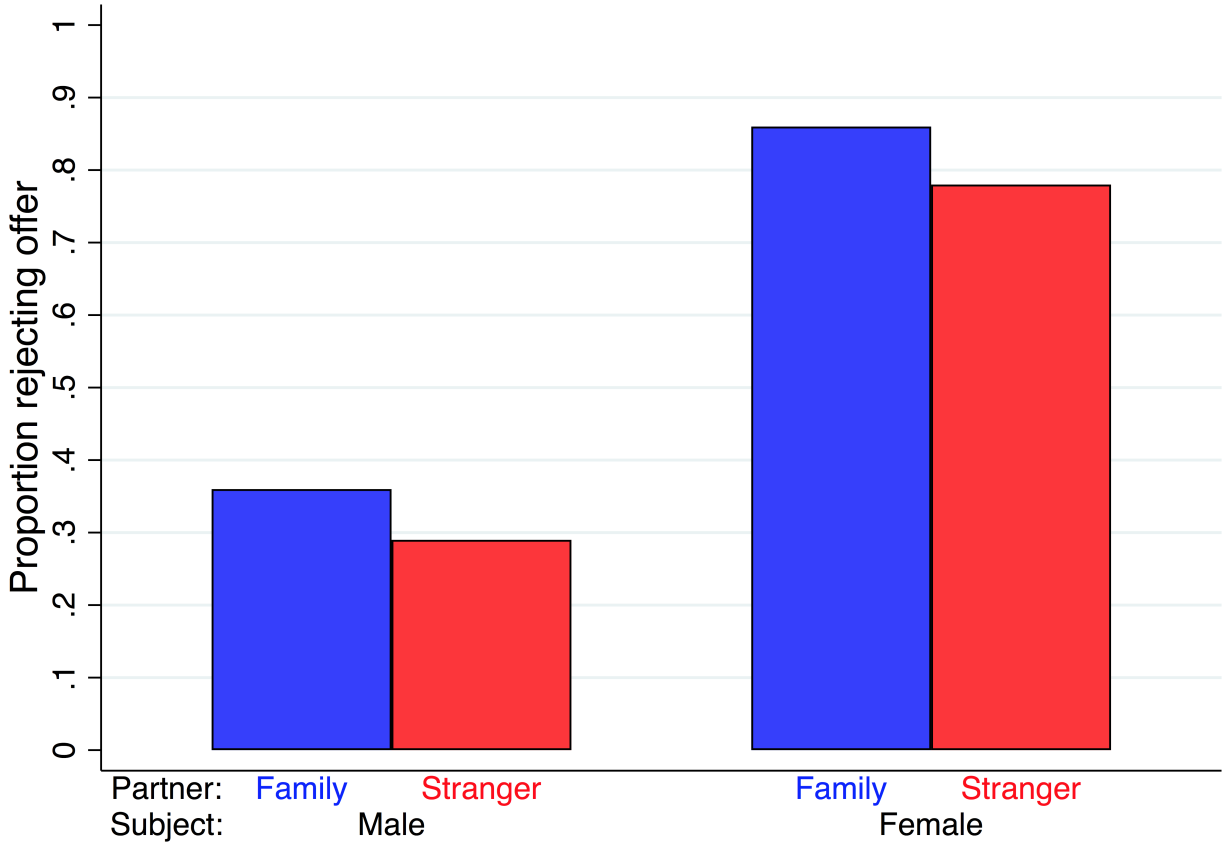
Table A2: Marginal effect of information on choice, conditional on guessing the partner's top rank

Dependent variable	Probability of choosing the partner's preferred option			
Sub-sample	Initial guess: wrong	Initial guess: right	Initial guess: wrong	Initial guess: right
Experiment	LAB1		LAB2	
	(1)	(2)	(3)	(4)
Marginal effects				
Male - Family	55.6%***	24.1%*	27.4%***	0.0%
	(10.3%)	(13.9%)	(6.6%)	(3.0%)
Male - Stranger Opposite sex	-1.7%	-19.4%	15.4%	-1.2%
	(12.2%)	(17.7%)	(9.7%)	(6.2%)
Male - Stranger Same sex			4.3%	1.8%
			(10.7%)	(4.5%)
Female - Family	4.2%	-21.1%	4.5%	-3.8%
	(15.1%)	(14.3%)	(6.0%)	(3.8%)
Female - Stranger Opposite sex	-0.4%	-1.4%	25.0%**	1.6%
	(13.7%)	(16.4%)	(10.6%)	(6.9%)
Female - Stranger Same sex			6.1%	-2.4%
			(10.0%)	(5.1%)
Overall	14.2%**	-5.4%	13.7%***	-1.1%
	(7.0%)	(8.4%)	(3.5%)	(1.9%)
No information	27.1%	75.3%	24.3%	87.0%
Information	41.4%	69.9%	38.1%	85.9%
N. Observations	204	132	803	1629

Standard errors (in percentage points) are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

Marginal effects are estimated using a regression of the dependent variable on a fully interacted set of an informed treatment dummy, a gender dummy, and dummies for each matching type. The regression is estimated separately for subjects who incorrectly guessed the top rank of their partner, and subjects who guessed correctly. The marginal effects of the information treatment that are reported in the Table are obtained using the 'margins, dydx' command in Stata. In LAB2 observations for food and drinks are combined. This means that in LAB2 there are up to four observations per subject: food and drink for family matching; and food and drink for stranger matching – except for subjects matched with a computer, in which case there are two observations per subjects: food and drink for family matching.

Figure A4: Any rejection in the ultimatum game



Notes: Figure A4 reports sample proportions of subjects rejecting any of the offers made in the ultimatum game. This Figure displays results when we restrict the sample to LAB1 subjects who show sufficient understanding of the ultimatum game. That is, if they can answer three test questions on Player 1 and Player 2 earnings correctly. Similar results are obtained using all subjects. Since this game was played with a strategy method, we observe willingness to reject for all subjects, irrespective of the choice made by the proposer. Similar results are found if we restrict the attention to rejections of offers below 60% instead. The differences between Male and Female subjects are both statistically significant.

Table A3: Testing demand for ex ante instrumental agency, restricting sample to mother-son pairs in LAB1

Dependent variable	Ex-ante agency: paying to get own pick before knowing partner's pick for self
Experiment	LAB1
	(1)
Expected gain from choosing own pick: Drink	-0.247
	(0.592)
Male - Stranger Opposite sex	0.275**
	(0.140)
Female - Family	0.260*
	(0.135)
Female - Stranger Opposite sex	0.025
	(0.067)
Intercept (Male - Family)	0.041
	(0.096)
N. Observations	50

Note: Standard errors are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports IV estimates of the effect of the expected gain from agency on willingness to pay for agency. In LAB1 willingness to pay for agency equals 1 if the subject prefers half a glass of their preferred to a coin toss. At the time of making the decision, subjects do not know what the partner has selected for them. Instruments include a fully interacted set of dummies for gender, informed treatment, matching type, and own top rank. The sample is restricted to mother-son pairs.

Table A4: Testing demand for ex ante instrumental agency among respondents who select their own top ranked items for consumption in LAB1

Dependent variable	Ex-ante agency: paying to get own pick before knowing partner's pick for self
Experiment	LAB1
	(1)
Expected gain from choosing own pick: Drink	-0.282
	(0.331)
Male - Stranger Opposite sex	-0.0461
	(0.078)
Female - Family	-0.020
	(0.086)
Female - Stranger Opposite sex	-0.119
	(0.076)
Intercept (Male - Family)	0.339***
	(0.090)
N. Observations	251

Note: Standard errors are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports IV estimates of the effect of the expected gain from agency on willingness to pay for agency. In LAB1 willingness to pay for agency equals 1 if the subject prefers half a glass of their preferred to a coin toss. At the time of making the decision, subjects do not know what the partner has selected for them. Instruments include a fully interacted set of dummies for gender, informed treatment, matching type, and own top rank. The sample is restricted to individuals with consistent choices, i.e. those who select their own top ranked item to consume.

Table A5: Testing demand for ex post pure agency, with control for whether respondents participated in the 'stranger' round before the 'family' round

Dependent variable	Ex-post agency: paying to get own pick after knowing partner's pick for self
Experiment	LAB2
	(1)
Actual gain from choosing own pick: Drink	-0.004
	(0.020)
Actual gain from choosing own pick: Food	0.078***
	(0.022)
Cost of getting own pick	-0.002***
	(0.000)
Male - Stranger Opposite sex	0.136***
	(0.044)
Male - Stranger Same sex	0.069*
	(0.037)
Male - Computer	0.023
	(0.039)
Female - Family	-0.017
	(0.022)
Female - Stranger Opposite sex	0.037
	(0.038)
Female - Stranger Same sex	-0.005
	(0.032)
Female - Computer	0.049
	(0.040)
Round order (stranger first)	0.031
	(0.019)
Intercept (Male - Family)	0.120***
	(0.025)
N. Observations	1432

Note: Standard errors are clustered at the pair level and in parentheses. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

The Table reports OLS estimates of the effect of the actual gain from agency on willingness to pay for agency. The dependent variables equals 1 if the subject forfeits a fraction of the showup fee to receive their top ranked food and drink, and 0 otherwise. At the time of making the decision, subjects know what the partner has selected for them. The cost of agency takes values 5, 20 and 50, equally randomized across subjects. 'Round order (stranger first)' is a binary variable that is 1 when the respondent participates in the a round matched with a stranger they participate in the round matched with family member.

Table A6: Husband's willingness to delegate in the field

	(1)
Husband's preference = choice he expects wife to make	0.010**
	(0.004)
Treated in the RCT	-0.000
	(0.005)
Intercept	0.985***
	(0.005)
N. Observations	1928

Standard errors in parenthesis. Stars correspond to significance levels: * for 10%, ** for 5%, and *** for 1%.

Notes: The Table reports OLS estimates of the effect of the actual loss from delegation on willingness to delegate. The dependent variables equals 1 if the subject is willing to delegate the choice of gift to his wife, and 0 otherwise. At the time of making the decision, subjects do not know what the partner has selected for them, but they have been asked to make a guess. The main regressor of interest equals 1 if the preference of the husband is equal to the choice he expects his wife to make.

Appendix B: Proxying the instrumental value of agency

By experimental design subject i receives their preferred bundle with probability 1 when exerting agency. In LAB1, a coin toss determines whether a subject receives their selected juice or the juice selected for them by their partner. Let d_i^k equal 1 if the partner selects i's preferred good k and 0 otherwise. The realized value of agency R_i for good k is thus:

$$R_i^k = \frac{1}{2}(1 - d_i^k)$$

where k only takes one value – juice. In LAB2, the subject receives the food and drink their partner chooses for them, unless they pay for agency. The value of agency is then:

$$R_i^k = 1 - d_i^k$$

for $k = \{\text{food,drink}\}$. When deciding whether to exert agency or not, subject i does not know d_i^k, but they can form an expectation of the probability that $d_i^k = 1$. Let this probability be written $E[d_i^k] = p_i^k$. The instrumental value of agency is thus the ex ante predicted value of R_i^k:

$$\begin{aligned} V_i^k &= E[R_i^k] \\ &= \frac{1}{2}(1 - p_i^k) \text{ in LAB1} \\ &= 1 - p_i^k \text{ in LAB2} \end{aligned}$$

As demonstrated earlier, p_i^k varies systematically with experimental treatments (e.g., matching type and information treatment) and characteristics of subject i (e.g., gender, preference ranking) – e.g., because they affect the partner's knowledge of i's preference ranking and their willingness to accommodate them. To capture this we estimate a regression model of the form:

$$d_i^k = \sum_m \sum_n \sum_r \sum_g D_m D_n D_{r_i^k} D_{g_i} + e_i^k \tag{8}$$

where the D_s's are categorical variables for matching type m, information dummy n, preference ranking of i for good k denoted r_i^k, and gender of i denoted g_i. Let $\widehat{p}_i^k$ denote the expected value of d_i^k predicted from the above estimated regression, estimated separately for LAB1 and LAB2 and for each good k. This value can then be used to obtain an estimate $\widehat{V}_i^k$ of V_i^k for each experiment and consumption good. Equivalently, we can regress model (8) directly on R_i^k to obtain an estimate of $E[R_i^k]$. In LAB1 $\widehat{V}_i^k$ varies between 8% and 44% (50% is the maximum feasible), with a mean of 25% (s.d. 9%). In LAB2 it varies between 8% and 100% for both food and drink, with an average of 36% (s.d. 24%) for food and 34% (s.d. 26%) for drink.

Ex ante demand for agency A_i^b can then be written as:

$$A_i^b = \alpha + \sum_k \beta_k R_i^k + \gamma C_i + \sum_m \sum_g D_m D_{g_i} + u_i \tag{9}$$

where C_i denotes experimental variation in the price of agency (only in LAB2) and where the realized value of agency R_i^k is instrumented using $\widehat{V}_i^k$. This is equivalent to using $\sum_m \sum_n \sum_r \sum_g D_m D_n D_{r_i} D_{g_i}$ as instruments in equation (9). Gender and matching type dummies are included as controls because they may affect agency directly – for instance because women are reluctant to impose their own choice when matched with a male family member.

We also estimate a version of model (9) for ex post agency A_i^p:

$$A_i^p = \alpha + \sum_k \beta_k R_i^k + \gamma C_i + \sum_m \sum_g D_m D_{g_i} + u_i$$

Here no instrumentation/prediction is required since, by construction, i knows R_i^k when deciding A_i^p.