Culture and the formation of gender-specific skills in an agrarian society *

Niels Kemper † Pia Naima Unte ‡

April 13, 2015

*We would like to express our gratitude to the Stiftungsfond of the Deutsche Bank for funding this field research and to Rainer Klump from the Goethe-University, Frankfurt, for assistance in writing the research proposal. We are also grateful to Tekie Alemu and Tadele Frede from Addis Ababa University for their intellectual and logistic support throughout the research project. Further, we are indebted to our field supervisor and to our enumerators for their diligent and thorough field work. We thank Markus Goldstein and other participants at the CSAE Conference 2012 at St Catherine's College, Oxford, as well as Agnes Quisumbing and other participants at the 2015 World Bank Conference on Land and Poverty for their helpful comments. We also acknowledge valuable comments of participants 2014 EEA-ESEM Conference, Toulouse, and the 2014 Conference on Development Economics, Passau. Further, we are grateful for helpful comments from Markus Frölich, Matthias Heinz, Heiner Schumacher and Albrecht Bohne.

†University of Mannheim, Department of Economics, L 7, 3–5, D–68131 Mannheim, ☎+49 (0)621 1805, ✉niels.kemper@uni-mannheim.de

‡University of Mannheim, Department of Economics, L 7, 3–5, D–68131 Mannheim, ☎+49 (0)621 1824, ✉pia.unte@gess.uni-mannheim.de

Culture and the formation of gender-specific skills in an agrarian society

Abstract

This study examines whether cultural norms arising from traditional agricultural practices affect the formation of gender-specific skills. We hypothesize that a culturally induced division of labor along gender lines generates gender-specific skills. As opposed to the traditional measurement of skills, which measures skill levels based on the type of tasks or abilities observed in certain occupations, we measure skills directly using a controlled field experiment in rural Ethiopia. Comparing women with exposure to the 'plow culture' with women without exposure to such cultural norms, and with men in general, we find a clear division of labor along gender lines between domestic and non-domestic work. We show that women exposed to the 'plow culture' are particularly skilled in exercising a light manual task resembling everyday work in the domestic sphere. Drawing on secondary data on the time-use of Ethiopian adults, we find supporting evidence that women with exposure to the 'plow culture' specialize in tasks from the domestic sphere. Thus, culturally-induced skill differences arguably are a neglected explanation for gender disparities in labor income.

Keywords: Skills, Gender Roles, Plow, Field Experiment, Ethiopia
JEL Classification: C93, J16, J24, N50

1 Introduction

Gender disparities in labor income are pervasive throughout the world.[1] Women are commonly over-represented in low-paying activities.[2] To explain gender disparities in labor income, economists have focused on gender-specific factors, such as differences in qualifications, as well as on the overall wage structure, i.e., the prices set for various labor market skills and the rents received for employment in particular sectors. It has been argued that female-dominated occupations require skills that generate lower returns than those required in male-dominated occupations (Blau and Kahn, 1992, 1996, 2000). Skills have been demonstrated to be a major determinant of returns to labor (Autor, 2014; Hanushek et al., 2013; Juhn et al., 1993), and they may be a major determinant of gender disparities in labor income if they differ by gender. Moreover, the influence of culture on economic outcomes has recently been emphasized (Guiso et al., 2006).[3] It has been suggested that gender gaps in wages and prospects for advancement may be attributed to men being culturally more inclined to competition than women (Gneezy et al., 2009). In this study, we examine whether culture affects the formation of gender-specific skills, hypothesizing that cultural persistence in a division of labor along gender lines generates gender-specific skills and, possibly, disparities in labor income.

Both culture and skills are somewhat elusive concepts and hard to operationalize for the research issue at hand. We build on work by Alesina et al. (2011, 2013)

[1] Ñopo et al. (2012) estimate a gender earnings gap ranging from 8 to 48 percent of average female earnings, with some regions, notably Sub-Saharan Africa, seeing more severe differences in earnings than others. For Kenya, for example, the gender wage gap is estimated to be 40 percent (Kabubo-Mariara, 2003), and for Botswana, 80 percent (Siphambe and Thokweng-Bakwena, 2001). The gender wage gap in hourly wages in Ethiopia's manufacturing sector is estimated to be 73 percent (Temesgen, 2006), the earnings gap in urban Ethiopia to be 78 percent (Appleton et al., 1999).

[2] Across countries, women are much more likely than men to be engaged in low-productivity activities, most commonly in the informal sector and in agricultural activities. If women are self-employed outside the agricultural sector, they tend to have small, informal, low-profit businesses, which they often operate from their homes. In the domestic sphere, women bear the main responsibility for housework and care while men are mainly responsible for market work. Relative to men, women devote one to three hours more per day to housework, two to ten times the amount of time per day to care, and one to four hours less per day to market activities (The World Bank, 2012). If women in rural Sub-Saharan Africa earn cash, their work is generally less remunerative than men's as they are largely restricted to activities from within the domestic sphere. As income-earning activities they sell, for example, prepared snacks, beer, and soap, or offer services such as hair plaiting (Bryceson, 2002).

[3] Culture is defined as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation."(Guiso et al., 2006, p. 23)

and Boserup (1970), who attribute existing cross-cultural differences in norms regarding the role of women in society to the usage of the plow in pre-industrial agriculture. Societies that introduced the plow to historic agriculture demonstrate less equal gender norms today than those that did not introduce the plow. This is reflected in less female participation in the labor force, entrepreneurial activities, and politics. Societies that introduced the plow also have lower fertility rates than those that did not. In plow agriculture, children and women are relatively less useful in the field than men, because operating the plow requires significant physical strength. Thus, men tend to specialize in field work while women specialize in home production and other work within the domestic sphere. In the following, we refer to such social norms as the 'plow culture.'

We hypothesize that if norms about a division of household labor along gender lines persist, then women with exposure to the plow culture are more likely to develop skills in tasks that resemble the activities regularly exercised in the household. Such tasks are typically non-marketable in the sense that they do not generate monetary returns outside the household. To investigate this in isolation, we designed a controlled field experiment conducted in rural Ethiopia. The outcome variable of interest is experimentally generated output from incentivized work. The experimental setup features a simple, manual task. The important feature of this task is that it can be considered to be context neutral for men in general and to some extent for women without exposure to the plow culture, while it strongly resembles tasks predominantly exercised by women with exposure to the plow culture. We argue that the task captures both effort and skills for women exposed to the plow culture, while it captures only effort for all other experimental subjects.[4] In the empirical analysis, we exploit variation in gender and in exposure to the plow culture—as proxied by the persistent cultivation patterns of crops favorable to the introduction of the plow to historic agriculture[5]—to disentangle the effects of effort and skill on our experimental outcome.

The analysis in this paper is based on data collected during experimental ses-

[4]Becker (1985) implicitly defines the term 'effort' as the amount of physical and mental energy invested in work. It is different from the set of skills and responsibilities required on a specific job, and from the particular set of skills and experiences of a specific worker. The term skill is commonly defined as "the ability to do something that comes from training, experience, or practice" (Merriam Webster, 2014).

[5]Only certain types of crops were favorable to the introduction of the plow to pre-industrial agriculture, which follows a marked geographic pattern related to the ecology (Pryor, 1985). Ecological factors determining this geographical pattern are constant over time (Wet, 1977; Ehret, 1979). See Section 3.1 for details.

sions à 20 participants held in 25 geographically dispersed *kebeles* (peasant associations) with randomly selected household heads from agricultural households in rural Ethiopia.[6] Using a field experiment is particularly useful for the research issue at hand since (i) skill and effort are difficult to observe, and even more difficult to disentangle, in traditional survey data, (ii) it allows us to choose the most appropriate experimental task for the research issue at hand, and (iii) it can be conducted in the naturally occurring environment of the experimental subjects. In addition to the experimental data, we draw on secondary data from the 1997 round of the Ethiopian Rural Household Survey (International Food Policy Research Institute, 2011), as it contains detailed data on time-use in Ethiopian households and allows us to cross-check the findings from the experimental data.[7]

Drawing on variation in the household-level experimental and survey data within a country rather than variation across countries limits the generalizability of our findings to other settings, but it allows us to examine cultural transmission channels. Cultural norms, as reflected in attitudes, beliefs, and preferences, have been shown to be influenced by one's family as well as the local environment extending beyond one's family (Fernández, 2007, 2011, 2013). We consider the household and the *kebele* as two potential transmission channels for the plow culture, and hypothesize that the household is the primary transmission channel since the decision regarding the division of labor along gender lines is made within the household.

In line with Alesina et al. (2011, 2013) and Boserup (1970), we find a significant gender-specific effect of being exposed to the plow culture on the probability of working primarily in the household. We also find that women with exposure to the plow culture clearly devote more time to activities in the domestic sphere than women without exposure to such cultural norms or than men in general. Domestic tasks include caring for children, cleaning, fetching water, gathering firewood, and preparing food and coffee, and making dungcakes. The gender-specific responsibility for these tasks is particularly pronounced when exposed to the plow culture. Further, we find that women who are exposed to the plow culture devote less time to activities outside the home, such as tending livestock, farming, and the marketing of household produce, than women without exposure to such

[6] *Kebeles*/peasant associations are the smallest administrative unit of Ethiopia.

[7] The 1997 round of the Ethiopian Rural Household Survey (ERHS) was undertaken by the Department of Economics of Addis Ababa University in collaboration with the International Food Policy Research Institute and the Centre for the Study of African Economies of Oxford University.

norms or than men in general. In our field experiment, we observe that women with exposure to the plow culture clearly outperform both men and the women without such exposure. We attribute this to the additional time spent on domestic tasks by women with exposure to the plow culture. Domestic work, unlike seasonal agricultural work, involves a frequent repetition of tasks resembling our experimental task. The more time is spent on sorting and picking processes in the domestic sphere, the more skilled are the participants in performing the experimental task. With regard to the transmission of the plow culture, we find that the corresponding social norms are mainly transmitted through one's family.

Among practitioners, it is widely recognized that the lack of marketable skills is a major obstacle to gender equality in the labor markets in developing countries (The World Bank, 2012). This may even be more true for agrarian economies, in which the natural conditions and gender roles cause women to overspecialize in tasks for which no market exists. As opposed to widespread markets bundling land and labor in rental and sharecropping contracts, by and large a male domain, as far as we know, no markets for household chores have historically emerged in agrarian economies. Consequently, women contribute less to household income, which is commonly associated with severe welfare consequences. Income in the hands of a mother improves child health and increases the share of household spending on nutrients, health, education, and housing, as compared to income controlled by men (Thomas, 1990, 1994). The effect on health, nutrition, and education may be limited to (older) girls, though, if women have a preference for spending resources on girls rather than boys (Duflo, 2003; Rangel, 2006).

Would a transition from an agrarian to an early industrial society close the gender gap in marketable skills and reduce gender inequality, especially for women exposed to the plow culture?[8] If history provides a lesson, the answer is partly yes. Goldin and Sokoloff (1984) argue that agricultural areas with low marginal productivity of women (and children) relative to men are favorable to industrialization. The lower the relative productivity of females in the pre-industrial agricultural or traditional economy, the earlier the manufacturing sector evolves and the more the relative wage for women increases. Goldin and Sokoloff (1982) show that early industrialization in the American Northeast experienced both a

[8]The government of Ethiopia sees the industrial sector as a key strategic pillar for achieving gender equality under the current "Growth and Transformation Plan," a national five-year plan bundling public action and assistance from international donors (Ethiopian Agricultural Transformation Agency , 2014).

dramatic increase of the female labor force in the manufacturing sector and a catch-up in female wages relative to male wages. However, women accumulated skills on the job and received less job training than men. Hence, the narrowing of the wage gap was not due to the narrowing of the skill gap, but to the substitution (associated with the mechanization and the separation of tasks in the new production processes) of cheaper, unskilled, female labor for more expensive, skilled, male labor.

As far as we know, the present paper is the first study to examine how culture shapes skills. It measures skills directly through an experimental approach. We disentangle whether cultural transmission works through the family or the local environment. Furthermore, we establish a link between the vibrant, but as yet separated, academic debates around gender disparities and the role of skills in labor market outcomes. Finally, we contribute to deepening the understanding of how traditional agricultural practices affect contemporary socio-economic outcomes.[9]

The remainder of this paper is organized as follows. We begin, in Section 2.1, by developing the hypotheses tested in the paper and, in Section 2.2, by describing the experimental design used to test them. In Section 2.3, we describe the data used in the empirical analysis and draw on the 2007 Population and Housing Census (Central Statistical Agency, 2007a) to confirm that our experimental data is representative of the underlying population in terms of its observable characteristics. In Section 3.1, we present our empirical strategy before we present estimates of the relation between the plow culture and the division of labor along gender lines in Section 3.2. We proceed with presenting estimates of the relation between the plow culture and gender-specific skills in Section 3.3. In Section 3.4, we present estimates exploring the primary transmission channel of the plow culture. We conclude in Section 4.

[9]Talhelm et al. (2014) report a persistent effect of a history of farming rice in southern China and wheat in northern China on psychological differences within China today. The need for cooperation and coordination to build and use irrigation systems in rice agriculture in the past is linked to cultural interdependence and holistic thinking in southern parts of China today. Moreover, a higher functional interdependence between farmers, as compared to herders, makes cultures that traditionally specialized in farming collectivistic, and cultures that traditionally specialized in herding individualistic (Nisbett et al., 2001).

2 Experimental Design

2.1 Hypotheses

This research sets out to examine the link between culture and the formation of skills along gender lines. We draw on variation in gender and exposure to social norms regarding gender roles, proxied by persistent patterns of crop cultivation, to show how these norms affect the formation of gender-specific skills. We test three hypotheses in this paper.

First, social norms regarding the role of women in society have been shown to be associated with historic agricultural practices. In contrast to shifting cultivation, plow agriculture requires significant physical strength and gives men an advantage in field work relative to women. Hence, in societies using the plow in pre-industrial agriculture, men tended to specialize in field work while women focused on home production and other work within the domestic sphere. This division of labor along gender lines generated persistent norms about gender roles in society, with the dominant belief that the appropriate role for women is within the domestic sphere (Alesina et al., 2011, 2013; Boserup, 1970). *Thus, women with exposure to the plow culture are more likely to devote most of their labor to work within the domestic sphere than either women without exposure to such cultural norms or than men in general.*

Second, if social norms and beliefs about a division of labor along gender lines persist, both inside and outside the household, women with exposure to the plow culture will differ systematically from men and from women without such exposure with respect to the tasks they perform regularly and hence the skills they possess. *Women with exposure to the plow culture are more likely to develop skills in tasks resembling the activities regularly exercised in the household than are women without exposure to such cultural norms or than are men in general.*

This hypothesis is straightforwardly formulated, albeit less straightforward to test given that skills are difficult to observe.[10] Traditionally, skill levels are measured by the type of observed tasks or abilities workers demonstrate in certain occupations. Keesing (1965, 1966), for example, postulates that labor skills determine the trade pattern for a broad group of manufactured products. Skills are divided into five observable categories (professional, technical, and manage-

[10]Individual skills, knowledge, and willingness to work are private, not public knowledge. Only their consequences are publicly observable (Smith, 1982).

rial; craftsmen and foremen—skilled manual workers; clerical, sales, and service; operatives—semi-skilled; laborers—unskilled) to calculate the skill requirements for given levels of exports and imports. To give another example, Autor et al. (2003) argue that the rapid adoption of computer technology changes the tasks performed by workers at their jobs, shifting the demand for human skills from routine to non-routine tasks. Non-routine manual activity is measured by the extent to which occupations involve observable activities such as managerial and interpersonal tasks, the ability to reason quantitatively, or the adaptability to work requiring set limits, tolerances, and standards, and eye–hand–foot coordination. Routine manual activity is measured by finger dexterity.

Recently, the relation between skills and wages has been analyzed using international data with skills being measured by the Programme for the International Assessment of Adult Competencies (PIAAC), an internationally harmonized test of cognitive and workplace skills measured by literacy, numeracy, and problem solving ability. This is in contrast to the bulk of the empirical literature on returns to education, which measures skill differences in terms of educational attainment. A large wage premium to cognitive skills is found. Across 22 countries, an increase by one standard deviation in numeracy is associated with an average 18 percent wage increase among prime-age workers (Autor, 2014; Hanushek et al., 2013).

While skill measurement in these studies clearly focuses on advanced economies with sophisticated production processes, it is of limited help for understanding the relation between skills and economic outcomes at the other end of the world income distribution, where skill formation by and large is not a matter of systematic training, but a learning-by-doing process. More importantly, being skilled in a task is a necessary but not a sufficient condition for being productive in a task, which obscures the effect of skills on labor market outcomes if skills are measured by the type of observed task and ability. Thus, we pursue an experimental approach which allows us to distinguish skill from effort in exercising a task, at least for a subsample of the potential labor force. We experimentally relate output and skills in a field lab experiment and measure skills directly rather than drawing on observable tasks alone. The experimental output was generated by incentivizing subjects to generate the experimental output (if there are binding individual participation constraints, participants do not exercise effort if they are not incentivized). To incentivize the participants, we used the contractual relations com-

monly found in land and labor markets in rural Ethiopia. We hypothesize that differences in experimentally generated output exist due to historically developed skill differences between men and women.

Third, social norms regarding a division of labor along gender lines and the natural place of men and women in society may be transmitted through various channels. Evidence exists that attitudes and preferences are transmitted from parents to children (Fernández, 2011). In the case of women's labor market participation, a mother's attitude toward working women has been found to be not only correlated with her daughter's, but also with her daughter-in-law's labor market decisions (Farré and Vella, 2013).[11] Moreover, individuals' beliefs and preferences are also influenced by the relationships and institutions of the local environment extending beyond one's family (Fernández, 2007, 2011, 2013).

We regard the household and the *kebele* as the two relevant channels for the cultural transmission of gender related social norms in rural Ethiopia. The decision regarding the division of labor along gender lines is made within the household, and such decisions have been shown to be related to one's family's culture (Farré and Vella, 2013; Fernández, 2013). *Hence, we expect exposure to the plow culture on the household level to be the primary channel for the transmission of the corresponding social norms.*

2.2 Experimental Procedure and Task

The analysis in this paper is based on data collected in 25 sessions at geographically dispersed study sites in rural Ethiopia. The experiment was conducted in 25 *kebeles* (peasant associations) in the two economically most active and most densely populated regions of Ethiopia (Amhara and Oromia).[12] Each session consisted of 20 household heads of randomly selected households.[13] The households to be invited were selected from land registries kept by *kebele* officials, which are widely believed to be up to date, as claims to land are registered there.[14] The

[11] With regard to the transmission of social norms to children-in-law, note that newly married couples in Ethiopia commonly move in with the husband's or wife's family. The wife's mother or mother-in-law will instruct her about homemaking during this time (Duncan and Hayden, 2008; Giel and Luijk, 1968; Hogan et al., 1999).

[12] See Figure 1 in Appendix E for a map of Ethiopia.

[13] For details on the sampling procedure, see Appendix B.

[14] *Kebeles* are the smallest administrative unit of Ethiopia and were set up after the 1974 revolution in order to implement land reforms. All land is owned by the government and households have to register with the *kebele* in order to obtain land (Dercon and Hoddinott, 2011).

invitation of participants was facilitated by one enumerator's arriving one day before the rest of the team to invite the participants. The enumerators were students from Addis Ababa University. Upon invitation, the participants were told that they could earn an unspecified amount of money during a research visit from a team from Addis Ababa University.

On the day of the experiment, each participant received a 20 Birr show-up fee (roughly 1.20 USD at the time of data collection). They were informed that they could quit the experiment at any time if they felt uncomfortable with any aspect of it. However, all participants chose to accomplish the task and the survey. The enumerators carefully explained the experimental task. Participants were also told how they would be paid. They answered control questions to demonstrate that they had understood the task and the payment procedure. The experimental task consisted of light manual work, i.e., a sorting and picking process. The basic means for conducting the experiment were two buckets, beans of three different colors (white, yellow and red), and a scale (see Figures 2 and 3 in Appendix E for pictures of the beans and the equipment used to conduct the experiment). Each participant received two buckets:

- A flat, wide bucket (blue) with a diameter of 22.1 cm. At the beginning of the experiment, this bucket was filled with beans of the three different colors; each bucket contained 700 g of beans of each color.

- A raised, narrow bucket (red) with a diameter of 16.3 cm. At the beginning of the experiment, this bucket contained no beans.

All flat and all raised buckets, respectively, had the same size, same shape, and contained exactly the same weights of red, yellow and white beans. For exercising the task, the participants picked beans of a certain color from the blue bucket and put them into the red bucket.[15] Participants were given 15 minutes for the sorting task. After the time was up, the enumerators collected the red buckets and weighted their contents on a scale. The achieved weight in grams is our measure for individual experimental output. Appendix C presents the details of how we established independence, i.e., measured individual rather than group effects, in the individual experimental outcomes.

[15] Yellow and white beans were virtually alike in size, but white beans weighed a little more. The red beans were bigger and heavier than the yellow and white beans. Therefore, the participants were not asked to pick out red beans. One-half of the sessions was assigned the task of picking out white beans, and the other one-half was assigned the task of picking out yellow beans.

We interpret the experimental task in terms of the responsibilities women commonly undertake in societies in Sub-Saharan Africa. While women predominantly supply the majority of labor inputs for food production and processing, as well as for household chores including care work, men divide their time mainly between farm work and leisure, providing minimal assistance to women in domestic work—see Arora and Rada (2013) for Ethiopia, Evers and Walters (2001) for Uganda, Pitcher (1996) for Mozambique, and Sow (2010) for Senegal. While the experimental task resembles the sorting and picking processes typical of agricultural production in Ethiopia, it also resembles the sorting and picking processes that are found in the domestic sphere, such as preparing food and gathering firewood. In Ethiopia, most of the women's time is spent in household reproduction activities like these (Frank, 1999). One important distinction between the sorting and picking processes in agriculture and the domestic sphere is that the former occur only seasonally, once or twice a year, while the latter are regular, even daily, activities. Therefore, we consider the experimental task as context-free for experimental participants who mainly engage in agricultural activities, while we consider it contextual for experimental participants mainly engaged in household activities.

More generally, research across disciplines suggests that task-specific skills depend, on the one hand, on the number of task repetitions (Brashers-Krug et al., 1996; Gilbert et al., 2001; Hauptmann and Karni, 2002; Karni et al., 1998; Newell et al., 2001; Salmon and Butters, 1995), and, on the other hand, on the time elapsed after the task was exercised (due to a process of memory consolidation) (Fischer et al., 2002; Karni, 1996; Karni et al., 1998). For our experimental task, we argue that the more time women spend on domestic work, which involves more frequent repetitions of simple sorting and picking processes than seasonal agricultural work, the more skilled they will be in exercising the experimental task.

In the context of agriculture, there are four cases in the decision on how to contract labor, land, and the claim to yields from land: wage labor, sharecropping, fixed-rent land leases, and ownership cultivation. The data on experimental output was hence generated under these four different compensation schemes. These varied across sessions but were the same for all participants within each particular session. The pay-off depended on the type of contract randomly assigned to each session and the experimental output achieved (except for the wage contract,

where the pay-off was fixed and independent of individual performance). These contracts were randomly assigned to six sessions each. The wage contract was assigned to seven sessions. See Appendix D for further details on the compensation schemes.

2.3 Data

We use data from three different data sources: First, and most importantly, we use data from the field lab experiment collected together with some survey data on the participants. We jointly refer to these as the experimental data. Second, we draw on the 1997 round of the Ethiopian Rural Household Survey (ERHS), as this round provides information on time-use for the different activities of men and women in rural Ethiopia: it contains a survey instrument concerning the duration of various activities performed by household heads and co-residing spouses during the previous day. We refer to these as the survey data. Finally, we use the 2007 Population and Housing Census of Ethiopia, the most recent census wave available, to check for the representativeness of the experimental sample for the underlying population. We refer to these as census data. For all the data sources, see Appendix A for the definitions of all the variables used in the analysis.

For the empirical analysis we use both experimental and survey data for two reasons. These data sources complement each other in the sense that the experimental data contain measures for the experimentally generated output and the survey data do not, while the survey data contain comprehensive data on time-use, which the experimental data do not. Furthermore, there is a significant overlap of information collected in both data sets which allows us to cross-check our findings in a number of cases. However, there is a major shortcoming to this approach. In the experimental data, the household head is defined to be the person with the best knowledge of household decision making and household resources.[16] Unfortunately, we were not able to find any definition of the household head in the ERHS documentation.[17] Potentially different definitions of the head may limit the comparability of findings across data sets. While the first definition is very specific and uses decision making—a rather male domain in rural Ethiopia—as

[16]Note that female household heads do not necessarily live by themselves. Only 4 percent of the women in the experimental sample reported living by themselves.

[17]In the census data, the head is defined as "[a]ny member of the household who is recognized as the head by the other member of the household" (Central Statistical Agency, 2007b, p. 29).

the criterion, the absence of a clear definition in the survey data leaves ample room for interpretation. Consequently, not necessarily the same type of person may be identified as the household head in the two data sets. We attempt to mitigate this problem by comparing our experimental data with survey data for both a sample containing only household heads and a sample containing household heads and their co-residing spouses. We restrict both survey samples to persons in the Amhara and Oromia regions aged 20 to 80, i.e., to persons in the same age group as the experimental subjects.[18] Table 1 provides descriptive statistics for the variables, from both the experimental and the survey data, that are used in the empirical analysis.

Furthermore, we compare the experimental data with census data to check whether they are representative for the Amhara and Oromia regions in terms of observable characteristics. The more similar are the averages of the observables from the experimental survey and the census, the more representative are the findings from the field lab for the underlying population. Table 2 presents descriptive statistics for the observable variables in both the experimental survey and the population census, namely, gender, age, age square, marriage status, family size, and literacy. As with the survey data, we restrict data taken from the census to persons aged 20 to 80 in the regions of Amhara and Oromia and look at both a sample consisting of only household heads and at a sample consisting of household heads and their spouses.

The means from the experimental sample are quite similar to those of the 'heads only' census data with respect to age and age squared. In terms of marriage status and family size, however, the sample of experimental subjects appears to be more similar to the 'household heads and spouses' census data. Two differences between the experimental subjects and the census are noteworthy. First, the share of women in the experimental sample is clearly lower than the share of females in the census data. We presume that this due to the differences in the definition of the household head. Second, individuals in the experimental survey appear to be clearly more literate than individuals in both census samples. We presume that this is due to differences in collecting information on literacy.

[18] After cleaning the ERHS data for inconsistencies (some men, for example, reported breastfeeding as an activity performed during the previous day) and missing values, we are left with 2,194 household heads and spouses (1,063 men and 1,131 women). Further restricting the sample to individuals aged 20 to 80 in the regions of Amhara and Oromia leaves 870 individuals (395 men and 475 women) out of which 533 are household heads (386 men and 147 women).

While the experimental survey classifies an individual as literate if the individual has ever received schooling (irrespective of the length and intensity), the census distinguishes between literacy and schooling. The latter probably gives a more accurate measure of literacy, implying that we may be overestimating the true literacy rates in the experimental sample.

Finally, we would like to discuss an interesting consequence of our definition of the household head. When the facilitator visited the sampled households for the invitation to the experimental session, he asked for the person with the best knowledge on household decision making and household resources to attend the meeting. Comparing households with exposure to the plow culture to those not exposed to such social norms, this definition led to dramatically different probabilities that the household would send a female representative to the session the next day. While 1 out of 15 households among the former would send a female representative, 1 out of 4 households from among the latter would send a female representative. We interpret this as prima facie supporting evidence for our first hypothesis, as women not exposed to the plow culture seem to be more engaged in household decision making and managing household resources (possibly contributing own-income) and less constrained in participating in activities outside the household (including, obviously, research visits from Addis Ababa University).

3 Empirical Strategy and Results

In this section we present and discuss the results from reduced form least squares estimations. First, we present our empirical strategy. Then, we establish the relation between culture and the division of labor along gender lines as well as the relation between social norms and the development of gender-specific skills. Finally, we present evidence that the household rather than the *kebele* is the main channel of cultural transmission.

3.1 Empirical strategy

The key explanatory variable in our empirical strategy is the type of crop which a household cultivates (or interactions therewith). Hence, we start by explaining the channel through which the type of crop cultivated by a household affects social

norms regarding the division of labor along gender lines and, hence, the formation of gender-specific skills:

First, the cultivation of crops in Ethiopia follows marked geographic patterns related to the ecology. Certain crops have been cultivated since 13,000 B.C., according to some estimates, and their cultivation has followed a distinct and persistent geographic pattern ever since people started to intensively consume and domesticate grains, including wheat, barley, and teff, at the end of the Pleistocene period (Wet, 1977; Ehret, 1979).

Second, some crops were more favorable to the introduction of the plow to pre-industrial agriculture than others. The plow is more beneficial for crops that require large tracts of land to be prepared in a short period of time to produce a sufficient amount of food calories to support a family, and for crops that can only be grown in soils that are not shallow, not sloped, and not rocky. For these plow-positive crops, the plow increases productivity. Hence, the introduction of the plow to historic agriculture is likely when such crops are cultivated. These crops include teff, wheat, barley, and wet rice. Plow-negative crops, in contrast, require less land to be prepared over a long period of time in order to support a family, and/or can be cultivated on thin, sloped, or rocky soils, where using the plow is difficult. Using a plow does not increase productivity when such crops are cultivated, thus, the introduction of the plow to historic agriculture is unlikely. They include maize, millet, and root and tree crops (Pryor, 1985). It should be emphasized that the geographic pattern in which certain types of crops are cultivated is unrelated to the location and borders of the present-day administrative units.[19] Consequently, our data show that households cultivating either plow-positive or plow-negative crops typically reside within the borders of the same *kebele*.

Third, as opposed to shifting cultivation, the introduction of the plow to historic agriculture entailed a division of labor with men specializing in field work and women focusing on home production and other work within the domestic sphere. This division of labor along gender lines resulted from the physical strength required to use the plow, which gave men an advantage in field work and generated norms about gender roles in society with the dominant belief that the appropriate role for women is within the domestic sphere. These norms have been shown to be persistent up to today (Alesina et al., 2011, 2013; Boserup,

[19]In fact, the formation of a *kebele* follows a rather arbitrary, man-made rule. Each *kebele* consists of at least 500 households, i.e., 3,500 to 4,000 persons.

1970).

Fourth, given that social norms are persistently transported by the cultivation of plow-positive crops, we should see that women in households cultivating plow-positive crops, or *kebeles* with a large share of households cultivating plow-positive crops, devote more labor to housework as compared to women from households cultivating plow-negative crops and as compared to men in general. As mentioned before, administrative units are man-made and unrelated to the geographic patterns of crop cultivation. Thus, many of our sampled *kebeles* contain both households classified as plow-positive and as plow-negative. Consequently, we explore two possible cultural transmission channels for social norms: *kebele* and household.

Our first set of regressions draws on variation in the cultivation of plow-positive crops and gender and an interaction of the two to estimate the gender-specific effects of cultural norms in the experimental data. It presumes cultural transmission through the household. The linear regression model is specified as follows:

$$\begin{aligned} y_{ij}^{exp} &= \beta_0 + \beta_1 fem_i + \beta_2 p\_pos_i + \beta_3 (fem_i \times p\_pos_i) \\ &\quad + \beta_4 c1_j + \beta_5 c3_j + \beta_6 c4_j + X_i\theta + \varepsilon_{ij} \end{aligned} \tag{1}$$

where $y_{ij}^{exp} \in \{house_i, log\_grams_i\}$ captures the particular outcome for individual $i = 1, \ldots, 20$ from session $j = 1, \ldots, 25$ from the experimental data. $house_i$ is a binary indicator equal to one if the main occupation of individual i is housework. log_grams_i denotes the logarithm of the experimentally generated output of individual i.

fem_i is a binary variable equal to one if individual i is female, and p_pos_i is a binary variable equal to one if the household of individual i cultivates a plow-positive crop. We consider an individual's household to be cultivating plow-positive crops if at least one of the crops, white teff, other teff, barley, or wheat, is cultivated on at least one of the household's plots.[20] In our sample, plow-positive

[20]If households cultivate plow-positive crops, they tend to do so on most of their land. According to the experimental survey data, households that cultivate plow-positive crops do so on two-thirds of their total available land (mean = 0.66 (in percentages of area), s.d. = 0.26). The share further increases if we drop land left fallow. Furthermore, the share of land devoted to plow-positive crops is highly correlated within a *kebele* (i.e., a session) in our data. We compute an intraclass correlation coefficient of 0.30. In the empirical analysis using the experimental data, the standard errors are clustered at the session level.

crops account for one-half of the land under cultivation (white teff 20.3, black and mixed teff 12.4, barl ey 6.0 and wheat 11.6 percent). Presuming that social norms regarding the role of women are persistently transmitted through the cultivation of plow-positive crops, the binary indicator captures the effect of the plow culture on the outcomes of interest. Consequently, the estimated coefficient $\widehat{\beta}_3$ for the interaction $(fem_i \times p\_pos_i)$ is our main coefficient of interest.

$c1_j$, $c3_j$, and $c4_j$ are binary variables equal to one if a particular contract is assigned to session j (wage, fixed-rent or ownership contract). The sharecropping contract $c2$ is taken as the contract of reference. The contracts are randomly assigned and hence independent of the outcomes. In principle, they could hence be dropped from the regressions. We include them in the regressions with log_grams_i as dependent variable to make the data generating process transparent, and we include them in all other regressions based on the experimental data in order to have consistent specifications.[21] X_i controls for other factors characterizing individual i that may have an independent influence on y_{ij}^{exp}, including age, age squared, marital status, family size, literacy, membership in socio-political organizations, and per-capita expenditure. ε_{ij} is a random disturbance term for individual i in session j.[22]

Regressions as in Equation 1 are based on the assumption that the social norms of the plow culture are transmitted through one's household. However, one's preferences and beliefs are also influenced by the relationships and institutions of one's local environment, extending beyond one's family (Fernández, 2007, 2011, 2013). In addition to our measure of exposure to the plow culture at the household level, p_pos_i, we also construct a measure for exposure to the plow culture at the *kebele* level, $p\_pos\_keb_j$, to disentangle the household from the *kebele* as potential channels through which cultural norms are transmitted.[23] We consider a person to be exposed to the plow culture at the *kebele* level if the share of households in the *kebele* that cultivate plow-positive crops is greater than or equal to the average share of households in the sample that grow plow-positive crops (0.84 in the experimental sample). We extend the linear model described in

[21] Unsurprisingly, the estimated regression coefficients are insensitive to the exclusion of the contract variables.

[22] The empirical analysis would be closer to Alesina et al. (2011, 2013) if we used two-stage least squares and instrumented plow usage today. However, given that consistent two-stage least squares estimates require large samples, we restrict our analysis to reduced form estimates to avoid biased estimates using the small sample at hand (given $n = 49$ females in our sample).

[23] For the experimental data, the terms 'session' and '*kebele*' are used interchangeably.

Equation 1 and also estimate the following linear model:

$$\begin{aligned} y_{ij}^{exp} = {} & \gamma_0 + \gamma_1 fem_i + \gamma_2 p\_pos_i + \gamma_3 p\_pos\_keb_j + \gamma_4 (fem_i \times p\_pos_i) \\ & + \gamma_5 (fem_i \times p\_pos\_keb_j) + \gamma_6 (p\_pos_i \times p\_pos\_keb_j) \\ & + \gamma_7 (fem_i \times p\_pos_i \times p\_pos\_keb_j) + \gamma_8 c1_j + \gamma_9 c3_j + \gamma_{10} c4_j \\ & + X_i \lambda + \nu_{ij} \end{aligned} \tag{2}$$

where y_{ij}^{exp}, fem_i, p_pos_i, $c1_j$, $c3_j$, $c4_j$, and X_i are defined as above and $p\_pos\_keb_j$ is a binary indicator equal to one if the share of households in *kebele* j that cultivate plow-positive crops is greater than or equal to the average share of households in the sample that grow plow-positive crops. Consequently, the estimated coefficients $\widehat{\gamma_4}$ for the interaction $(fem_i \times p\_pos_i)$, $\widehat{\gamma_5}$ for the interaction $(fem_i \times p\_pos\_keb_j)$ and $\widehat{\gamma_7}$ for the triple interaction $(fem_i \times p\_pos_i \times p\_pos\_keb_j)$ are our main coefficients of interest.

In addition to the experimental data, we use data from the 1997 round of the ERHS for our empirical analysis. This round not only provides information on the time-use for the various activities performed by household heads and spouses during the previous day, but also contains information on the crops cultivated during the two main agricultural seasons, *belg* and *meher*. Hence, we can draw on the 1997 round of the ERHS to further support our evidence from the experimental data. Using the survey data we estimate the following linear regression model:

$$y_i^{ERHS} = \delta_0 + \delta_1 fem_i + \delta_2 p\_pos_i + \delta_3 (fem_i \times p\_pos_i) + Z_i \xi + \omega_i \tag{3}$$

where $y_i^{ERHS} \in \{house_i, time_{it}\}$ captures the particular outcome for individual i from the survey data. As in the experimental data, $house_i$ is a binary indicator equal to one if the main occupation of individual i is housework. $time_{it}$ denotes the time that individual i devotes to activity t, $t \in \{housework, childcare, livestock, farming, marketing, owntime\}$.

fem_i is a binary variable equal to one if individual i is female, and p_pos_i is a binary variable equal to one if individual i's household cultivates a plow-positive crop. Again, an individual's household is considered to be cultivating plow-positive crops if at least one of the crops teff, barley, or wheat is cultivated on at least one of the household's plots during at least one of the two agricultural seasons. In the ERHS sample containing heads and spouses (only heads), plow-

positive crops account for 46 (44) percent of the land under cultivation (teff 15 (14), barley 23 (22), and wheat 8 (8) percent).[24]

As in the experimental data, the binary indicator captures the effect of social norms regarding the role of women. Consequently, the estimated coefficient $\widehat{\delta}_3$ for the interaction $(fem_i \times p\_pos_i)$ is our main coefficient of interest. Z_i controls for other factors characterizing individual i that may have an independent influence on y_i^{ERHS}, including being a household head, family size, age, age squared, literacy, and per-capita expenditure. ω_i is a random disturbance term for individual i.

Again, we extend the linear model from Equation 3 to disentangle the household and the *kebele* as potential transmission channels of the cultural norms of the plow culture and also estimate the following linear regression model:

$$\begin{aligned} y_{ij}^{ERHS} = {} & \phi_0 + \phi_1 fem_i + \phi_2 p\_pos_i + \phi_3 p\_pos\_keb_j + \phi_4 (fem_i \times p\_pos_i) \\ & + \phi_5 (fem_i \times p\_pos\_keb_j) + \phi_6 (p\_pos_i \times p\_pos\_keb_j) \\ & + \phi_7 (fem_i \times p\_pos_i \times p\_pos\_keb_j) + Z_i \psi + \upsilon_{ij} \end{aligned} \tag{4}$$

where y_{ij}^{ERHS}, fem_i, p_pos_i, and Z_i are defined as above and $p\_pos\_keb_j$ is a binary indicator equal to one if the share of households in *kebele* j that cultivate plow-positive crops is greater than or equal to the average share of households in the sample that grow plow-positive crops (0.85 in the 'heads and spouses' ERHS sample, 0.81 in the 'heads only' ERHS sample). Consequently, the estimated coefficients $\widehat{\phi}_4$ for the interaction $(fem_i \times p\_pos_i)$, $\widehat{\phi}_5$ for the interaction $(fem_i \times p\_pos\_keb_j)$ and $\widehat{\phi}_7$ for the triple interaction $(fem_i \times p\_pos_i \times p\_pos\_keb_j)$ are our main coefficients of interest.

3.2 Culture and Division of Labor along Gender Lines

In this section we present empirical evidence for the influence of culture on the division of labor along gender lines. For both the ERHS and experimental data,

[24] If households cultivate plow positive crops, they tend to do so on most of their land. In both the ERHS samples of interest, households that cultivate plow-positive crops do so on more than one-half of their total available land (mean = 0.56 (in percentage of area), s.d. = 0.33). Furthermore, the share of land devoted to plow-positive crops is highly correlated within a *kebele*. We compute an intraclass correlation coefficient of 0.21 (0.22 in the sample containing only household heads). In the empirical analysis using the survey data, standard errors are clustered at the *kebele* level.

we reject the null hypothesis that men's and women's main occupation is housework with equal probability. For the ERHS data, this applies to both the 'heads and spouses' and the 'heads only' samples.[25] In addition, women with exposure to the plow culture are more likely to be working primarily in the household than women without exposure to such norms. For the ERHS data, again, this applies to both samples.[26] Using the regressions specified in Equations 1 and 3, we further explore the hypothesis that women with exposure to the plow culture are more likely to be primarily working in the household.

Table 3 (respectively, 4) presents the regression results from letting the effect of cultivating plow-positive crops vary by gender, i.e., the effect of being a female and exposed to the plow culture at the household level, on the probability of working primarily in the household using the experimental (respectively, ERHS) data. In Table 3, results from unweighted regressions are presented in columns (1) and (2); results from weighted regressions accounting for unequal selection probabilities into the random sample are presented in columns (3) and (4).[27] In the even-numbered columns, we present the results of regressions including further covariates. In Table 4, columns (1) and (2) show the results for the sample of household heads and spouses, while columns (3) and (4) present the results for the 'heads only' sample. Again, the even-numbered columns present regression results including further covariates.

Women clearly devote more time to housework than men. In Table 3, the probability of the main occupation being housework is between 57 and 64 percentage points higher for women than it is for men. In Table 4, this probability is even 92 to 97 percentage points higher for women than for men. Across specifications and datasets, these point estimates are highly significant. As hypothesized, we find a positive gender-specific increase of the effect of plow-positive crops on the probability of the main occupation being housework. In Table 3, all else equal,

[25] $H_0 : \overline{house}^{ERHS}_{fem=0} = \overline{house}^{ERHS}_{fem=1}$, $t = -80.6551$, $Pr(|T| > |t|) = 0.0000$, $n = 870$;
$H_0 : \overline{house}^{ERHS}_{fem=0,head=1} = \overline{house}^{ERHS}_{fem=1,head=1}$, $t = -45.4063$, $Pr(|T| > |t|) = 0.0000$, $n = 533$;
$H_0 : \overline{house}^{exp}_{fem=0} = \overline{house}^{exp}_{fem=1}$, $t = -20.6725$, $Pr(|T| > |t|) = 0.0000$, $n = 498$.

[26] $H_0 : \overline{house}^{ERHS}_{fem=1,p\_pos=0} = \overline{house}^{ERHS}_{fem=1,p\_pos=1}$, $t = -6.5853$, $Pr(|T| > |t|) = 0.0000$, $n = 474$;
$H_0 : \overline{house}^{ERHS}_{fem=1,p\_pos=0,head=1} = \overline{house}^{ERHS}_{fem=1,p\_pos=1,head=1}$, $t = -3.3008$, $Pr(|T| > |t|) = 0.0012$, $n = 146$;
$H_0 : \overline{house}^{exp}_{fem=1,p\_pos=0} = \overline{house}^{exp}_{fem=1,p\_pos=1}$, $t = -1.2334$, $Pr(|T| > |t|) = 0.2235$, $n = 49$.

[27] See Appendix B for details on the construction of the weights.

women exposed to the plow culture appear to have a higher probability, by 19 to 28 percentage points, to be working primarily in the household relative to women without such exposure. In Table 4, this gender-specific effect of the plow culture amounts to 13 to 19 percentage points. Across specifications and datasets, these point estimates are either significant or close to being significant at conventional significance levels.

3.3 Skills and Division of Labor along Gender Lines

We have demonstrated that women are more likely to primarily work in the household, particularly when exposed to the plow culture. We argue that this division of labor entails the formation of gender-specific skills. Women develop skills in simple manual tasks such as picking and sorting processes that, unlike in agricultural work, occur regularly (even daily) in domestic work. Domestic tasks include cleaning, fetching water, gathering firewood, and preparing food, coffee, and making dungcakes. Based on the common notion that skills develop with the number of repetitions of a task, we argue that, with respect to our experimental task, the more time women spend on domestic work, the more skilled they will be at exercising such tasks. To complete the evidence for the division of labor, we also present evidence for the time-use of adults in Ethiopian households. Then we present the findings from our controlled field experiment.

Following the classification of activities in the ERHS, we aggregate time-use to time devoted to housework, childcare, livestock tending, farming, marketing, and time for oneself.[28] Using these time-use categories as the outcome variables, we estimate the regressions as specified in Equation 3. In Tables 5 and 6 we present regression results when letting the effect of exposure to the plow culture on household level vary by gender for the sample containing heads and spouses (Table 5) and the sample containing only household heads (Table 6). In the even-numbered columns, we present the results when letting the effect of the plow culture vary by gender only. In the odd-numbered columns, we present the results of regressions including further covariates.

[28] In the ERHS, other employment/work is an additional category. We refrain from including this in our empirical analysis as only very few people reported pursuing an activity that falls into this category. A major limitation of the survey is that the survey instrument collecting information on time-use concerns only the previous day. Hence, we also refrain from including activities classified under non-employment activities away from home, as this category comprises extraordinary activities. See Appendix A for a definition of all the variables used in the empirical analysis.

As hypothesized, we find a significantly positive effect of being female on the amount of time devoted to housework. Women, on average, devote 4.7 to 5.3 hours per day more to domestic tasks than men do. Moreover, we find a positive gender-specific increase of the effect of plow-positive crops on the amount of time devoted to domestic tasks. All else equal, women exposed to the plow culture appear to devote an additional hour per day to domestic tasks relative to women without such exposure. This effect is significant across all specifications. We also find evidence that women exposed to the plow culture spend significantly more time on childcare than men or than women without exposure to such norms. Furthermore, we find that women exposed to the plow culture spend considerably less time on activities outside the home than women without such exposure. They devote less time to livestock tending, farming, or the marketing of household produce than men or than women without exposure to the plow culture. As a consequence of this culturally driven division of labor, we argue that women who are exposed to the plow culture are more skilled in our experimental task than are men or than are women who are not so exposed, as the experimental task resembles the tasks commonly undertaken in household work.

To test this hypothesis, we draw on the experimentally generated output from a simple manual task. Given that the experimental task resembles tasks more often undertaken by women with exposure to the plow culture than by women without such exposure, we expect that women exposed to the plow culture outperform men and outperform women who have not been exposed to such cultural norms. The task captures the effort and skills for women in households cultivating plow-positive crops while it captures only effort for men or for women in households cultivating plow-negative crops.

On average, the participants achieved an output of 141.1 g, with the average being 138.7 g for men (s.d. = 25.9 g) and 164.4 g for women (s.d. = 24.9 g). The difference between women with and without exposure to the plow culture amounts to 14.3 g. The null hypothesis that men and women, on average, achieve the same output is rejected.[29] More importantly, in the experimental task, women exposed to the plow culture significantly outperform women without exposure to such norms.[30]

In Table 7 we present the regression results letting the effect of social norms

[29] $H_0 : \overline{grams}^{exp}_{fem=0} = \overline{grams}^{exp}_{fem=1}$, $t = -6.6403$, $Pr(|T| > |t|) = 0.0000$, $n = 498$.

[30] $H_0 : \overline{grams}^{exp}_{fem=1,p\_pos=0} = \overline{grams}^{exp}_{fem=1,p\_pos=1}$, $t = -2.0548$, $Pr(|T| > |t|) = 0.0455$, $n = 49$.

on experimental output vary by gender. The results from unweighted regressions are shown in columns (1) to (3), the results from weighted regressions accounting for unequal selection probabilities in columns (4) to (6). In columns (1) and (4), we present the estimated gender effect on the experimental outcome controlling only for the underlying compensation scheme. In columns (2), (5), (3) and (6) we let the effect of the plow culture vary by gender and include further covariates.

As can be seen in Table 7, women significantly outperform men in the experimental task: the difference amounts to 16.1 to 21.5 percent. Moreover, we find a significant gender-specific effect, of around 14 to 15 percent, of cultivating plow-positive crops. We interpret the gender gap in the experimental outcome as revealing differences in skills driven by a culturally induced division of labor along gender lines.

3.4 Transmission Channels of Cultural Norms

Social norms regarding a division of labor along gender lines and the natural place of men and women in society may be transmitted through various channels. In addition to our measure of exposure to the plow culture at the household level, we also consider a measure for exposure to the plow culture at the *kebele* level, in order to disentangle the household from the *kebele* as two potential channels through which cultural norms may have been transmitted.

In Table 8 we present the regression results for regressing the probability of working primarily in the household on the triple interaction of gender, exposure to the plow culture at the household level, and such exposure at the *kebele* level, as specified in Equation 2. The results from the unweighted regressions are presented in columns (1) and (2); the results from weighted regressions accounting for unequal selection probabilities into the random sample are presented in columns (3) and (4). In the even-numbered columns, we present the results of regressions including further covariates.

Compared to women with no exposure to the plow culture, women with exposure to the plow culture only at the household level have a significantly higher probability, 40 to 45 percentage points higher, to be working primarily in the household. Moreover, comparing women with exposure to the plow culture at the household level, we find a significantly positive additional effect of exposure to the plow culture at the *kebele* level of between 62 and 66 percentage points. Exposure to the plow culture at the *kebele* level alone, however, does not appear

to have a significant effect on women's probability to be working mainly in the household. The household appears to be the main channel for the transmission of the social norms of the plow culture.

These findings are confirmed by the corresponding regressions (as specified in Equation 4) using the ERHS sample, which are presented in Table 9. In columns (1) and (2), we present the regression results for the sample of household heads and spouses, while the results in columns (3) and (4) are based on a sample of only household heads. In the even-numbered columns, we present the results of regressions including further covariates. Exposure to the plow culture only at the household level significantly increases women's probability to be mainly engaged in housework. Additional exposure to the plow culture at the *kebele* level has an additional positive effect on the probability of women to be primarily doing housework, which is also highly significant. We also find a significantly positive effect of women being exposed to the plow culture at the *kebele* level only.

Finally, Table 10 presents the regression results for regressing the experimental output on the triple interaction of gender, exposure to the plow culture at the household level, and such exposure at the *kebele* level. The results from the unweighted regressions are presented in columns (1) and (2), the results from weighted regressions accounting for unequal selection probabilities into the random sample are presented in columns (3) and (4). In the even-numbered columns, we present the results of regressions including further covariates.

We find that women with exposure to the plow culture only at the household level significantly outperform women without exposure to the plow culture: by 28 to 30 percent. Again, the triple interaction yields a significant additional effect of exposure to the plow culture at the *kebele* level, of 14 to 15 percent. Being exposed to the plow culture only at the *kebele* level does not have a significant effect on women's performance in the experimental task.

All in all, we interpret these findings as revealing that the social norms of the plow culture are mainly transmitted through one's family, which we capture through our measure of exposure to the plow culture at the household level.

4 Concluding Remarks

We hypothesize that cultural persistence in the division of labor along gender lines arising from traditional agricultural practices and transmitted primarily through

the household relates to the formation of gender-specific skills. We test this hypothesis combining evidence from a simple controlled experiment with survey data, especially on the time-use of Ethiopian households.

Our findings point toward a culturally driven division of labor. Men are found to devote their time to agricultural tasks, women with exposure to the plow culture appear to be restricted to low-return, year-round activities in the domestic sphere. Drawing on experimental data, we show that women who are exposed to the plow culture, which sees the natural place for women to be within the domestic sphere, generate more output in a simple manual task resembling low-return activities than women in areas where natural conditions did not favor the introduction of the plow to traditional agriculture. They also generate more output at this task than men in general. We attribute this fact to those women's being more skilled in this task due to a culturally induced division of labor along gender lines. With regard to the transmission of the plow culture, we find that the corresponding social norms are mainly transmitted through one's family.

References

Alesina, A., Giuliano, P., and Nunn, N. (2011). Fertility and the Plough. *The American Economic Review: Papers & Proceedings*, 101(3):499–503.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Appleton, S., Hoddinott, J., and Krishnan, P. (1999). The Gender Wage Gap in Three African Countries. *Economic Development and Cultural Change*, 47(2):289–312.

Arora, D. and Rada, C. (2013). Gender differences in time and resource allocation in rural households in Ethiopia. Paper prepared for the Annual Conference of the American Economic Association. Retrieved on 8 August 2014 from `http://www.aeaweb.org/aea/2014conference/program/retrieve.php?pdfid=421`.

Autor, D., Levy, F., and Murnane, R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843–851.

Becker, G. S. (1985). Human Capital, Effort, and the Sexual Division of Labor. *Journal of Labor Economics*, 3(1):S33–S58.

Blau, F. D. and Kahn, L. M. (1992). The Gender Earnings Gap: Learning from International Comparisons. *The American Economic Review: Papers and Proceedings*, 82(2):533–538.

Blau, F. D. and Kahn, L. M. (1996). Wage Strucutre and Gender Earnings Differentials: An International Comparison. *Economica*, 63(250):S29–S62.

Blau, F. D. and Kahn, L. M. (2000). Gender Differences in Pay. *The Journal of Economic Perspectives*, 14(4):75–99.

Boserup, E. (1970). *Woman's Role in Economic Development*. Earthscan, London.

Brashers-Krug, T., Shadmehr, R., and Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, 382:252–255.

Bryceson, D. F. (2002). The Scramble in Africa: Reorienting Rural Livelihoods. *World Development*, 30(5):725–739.

Central Statistical Agency (2007a). Population and Housing Census of 2007. Data Set, Addis Ababa.

Central Statistical Agency (2007b). The 2007 Population and Housing Census of Ethiopia: Enumerators' Instructional Manual. Retrieved on 8 August 2014 from `http://ecastats.uneca.org/aicmd/AfricaDirectoryEI/Documents/et2007ei_ethiopia_enumerator_manual.en.pdf`.

Dalet, D. (2014). Ethiopia: boundaries, regions, names. Retrieved on 15 January 2014 from `http://d-maps.com/carte.php?num_car=4258&lang=en`.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. John Hopkins University Press, Baltimore, MD.

Dercon, S. and Hoddinott, J. (2011). The Ethiopian Rural Household Surveys 1989–2009: Introduction. Retrieved on 3 June 2014 from `https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/244/download/4203`.

Duflo, E. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *World Bank Economic Review*, 17(1):1–25.

Duncan, A. and Hayden, M. (2008). Ethiopia Cultural Profile: Marriage. Retrieved on 9 October 2014 from https://ethnomed.org/culture/ethiopian/copy_of_ethiopian-cultural-profile#section-5.

Ehret, C. (1979). On the Antiquity of Agriculture in Ethiopia. *The Journal of African History*, 20(2):161–177.

Ethiopian Agricultural Transformation Agency (2014). National Growth & Transformation Plan. Retrieved on 8 December 2014 from http://www.ata.gov.et/priorities/national-growth-transformation-plan/.

Evers, B. and Walters, B. (2001). The Model of a Gender-Segregrated Low-Income Economy Reconsidered: Evidence from Uganda. *Review of Development Economics*, 5(1):76–88.

Falk, A. and Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1):39–58.

Farré, L. and Vella, F. (2013). The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labour Force Participation. *Economica*, 80(318):219–247.

Fernández, R. (2007). Alfred Marshall Lecture: Women, Work, and Culture. *Journal of the European Economic Association*, 5(2–3):305–332.

Fernández, R. (2011). Does Culture Matter? In Benhabib, J., Jackson, M. O., and Bisin, A., editors, *Handbook of Social Economics*, volume 1A. North-Holland, San Diego, CA.

Fernández, R. (2013). Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century. *The American Economic Review*, 103(1):472–500.

Fischer, S., Hallschmid, M., Elsner, A., and Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences of the United States of America*, 99(18):11987–11991.

Frank, E. (1999). Gender, Agricultural Development and Food Security in Amhara, Ethiopia: The Contested Identity of Women Farmers in Ethiopia. Mimeo, United States Agency for International Development, Washington D.C.

Giel, R. and Luijk, J. N. (1968). Patterns of Marriage in a Roadside Town in South-Western Ethiopia. *Journal of Ethiopian Studies*, 6(2):61–69.

Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The Neural Basis of Perceptual Learning. *Neuron*, 31:681–697.

Gneezy, U., Leonard, K. L., and List, J. (2009). Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society. *Econometrica*, 77(5):1637–1664.

Goldin, C. and Sokoloff, K. (1982). Women, Children, and Industrialization in the Early Republic: Evidence from the Manufacturing Censuses. *The Journal of Economic History*, 42(4):741–774.

Goldin, C. and Sokoloff, K. (1984). The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850. *The Quarterly Journal of Economics*, 99(3):461–487.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does Culture Affect Economic Outcomes? *The Journal of Economic Perspectives*, 20(2):23–48.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2013). Returns to Skills around the World: Evidence from PIAAC. NBER Working Paper 19762, National Bureau of Economic Research, Cambridge, MA.

Hauptmann, B. and Karni, A. (2002). Primed to learn: saturation of repetition priming and the induction of long-term memory. *Cognitive Brain Research*, 13(3):313–322.

Hogan, D. P., Berhanu, B., and Hailemariam, A. (1999). Household Organization, Women's Autonomy, and Contraceptive Behavior in Southern Ethiopia. *Studies in Family Planning*, 30(4):302–314.

International Food Policy Research Institute (2011). Ethiopia Rural Household Survey Dataset, 1989–2009. Retrieved on 3 June 2014 from `http://www.ifpri.org/dataset/ethiopian-rural-household-surveys-erhs`.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *The Journal of Political Economy*, 101(3):410–442.

Kabubo-Mariara, J. (2003). Wage determination and the gender wage gap in Kenya: Any evidence of gender discrimination? Research Paper 132, African Economic Research Consortium, Nairobi.

Karni, A. (1996). The acquisition of perceptual and motor skills: a memory system in the adult human cortex. *Cognitive Brain Research*, 5(1–2):39–48.

Karni, A., Meyer, Rey-Hipolito, C., Jezzard, P., Adams, M., Turner, R., and Ungerleider, L. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):861–868.

Keesing, D. B. (1965). Labor Skills and International Trade: Evaluating Many Trade Flows with a Single Measuring Device. *The Review of Economics and Statistics*, 47(3):287–294.

Keesing, D. B. (1966). Labor Skills and Comparative Advantage. *The American Economic Review*, 56(1/2):249–258.

Kemper, N. and Schumacher, H. (2014). Unfair Incentives: A Behavioral Note on Sharecropping. Mimeo, University of Mannheim, Mannheim.

Mas, A. and Moretti, E. (2009). Peers at Work. *The American Economic Review*, 99(1):112–145.

Merriam Webster (2014). skill. Retrieved on 3 June 2014 from `http://www.merriam-webster.com/dictionary/skill`.

Newell, K., Liu, Y., and Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychological Review*, 108(1):57–82.

Nisbett, R. E., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and Systems of Thought: Holistic Versus Analytic Cognition. *Psychological Review*, 108(2):291–310.

Ñopo, H., Daza, N., and Ramos, J. (2012). Gender earnings gaps around the world: a study of 64 countries. *International Journal of Manpower*, 33(5):464–513.

Pitcher, M. A. (1996). Conflict and Cooperation: Gendered Roles and Responsibilities within Cotton Households in Northern Mozambique. *African Studies Review*, 39(3):81–112.

Pryor, F. L. (1985). The Invention of the Plow. *Comparative Studies in Society and History*, 27(4):727–743.

Rangel, M. A. (2006). Alimony Rights and Intrahousehold Allocation of Resources: Evidence from Brazil. *The Economic Journal*, 116(513):627–658.

Salmon, D. and Butters, N. (1995). Neurobiology of skill and habit learning. *Current Opinion in Neurobiology*, 5(2):184–190.

Siphambe, H. K. and Thokweng-Bakwena, M. (2001). The Wage Gap between Men and Women in Botswana's Formal Labour Market. *Journal of African Economies*, 10(2):127–142.

Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *The American Economic Review*, 72(5):923–955.

Sow, F. D. (2010). *Intrahousehold resource allocation and well–being: The case of rural households in Senegal*. PhD thesis, University of Wageningen.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., and Kitayama, S. (2014). Large-Scale Psychological Differences Within China Explained by Rice Versus Wheat Agriculture. *Science*, 344(6184):603–608.

Temesgen, T. (2006). Decomposing Gender Wage Differentials in Urban Ethiopia: Evidence from Linked Employer-Employee (LEE) Manufacturing Survey Data. *Global Economic Review*, 35(1):43–66.

The World Bank (2012). Gender Equality and Development. World Development Report, The World Bank, Washington, D.C.

Thomas, D. (1990). Intra-Housheold Resource Allocation: An Inferential Approach. *The Journal of Human Resources*, 25(4):635–664.

Thomas, D. (1994). Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height. *The Journal of Human Resources*, 29(4):950–988.

Wet, J. (1977). Domestication of African Cereals. *African Economic History*, 3:15–32.

Tables

Table 1: **Observable characteristics (experimental and survey data)**

	Experimental Data	Survey Data	
		heads and spouses	heads only
outcome variables			
childcare	-	0.30	0.18
	-	(1.12)	(0.89)
	-	[870]	[533]
farming	-	2.06	2.85
	-	(3.49)	(3.94)
	-	[870]	[533]
grams	141.05	-	-
	(26.92)	-	-
	[500]	-	-
house	0.06	0.52	0.24
	(0.23)	(0.50)	(0.43)
	[498]	[870]	[533]
housework	-	3.30	1.78
	-	(3.73)	(3.17)
	-	[870]	[533]
livestock	-	1.00	1.29
	-	(1.94)	(2.20)
	-	[870]	[533]
marketing	-	0.37	0.40
	-	(1.41)	(1.49)
	-	[870]	[533]
owntime	-	3.61	3.85
	-	(2.29)	(2.24)
	-	[870]	[533]
treatment variables			
c1	0.28	-	-
	(0.45)	-	-
	[500]	-	-
c2	0.24	-	-
	(0.43)	-	-
	[500]	-	-
c3	0.24	-	-
	(0.43)	-	-
	[500]	-	-
c4	0.24	-	-
	(0.43)	-	-
	[500]	-	-

(. . .)

	Experimental Data	Survey Data	
		heads and spouses	heads only
female	0.10	0.55	0.28
	(0.30)	(0.50)	(0.45)
	[498]	[870]	[533]
head	-	0.61	-
	-	(0.49)	-
	-	[870]	-
plow_pos	0.84	0.85	0.81
	(0.37)	(0.36)	(0.39)
	[499]	[869]	[532]
plow_pos_keb	0.76	0.30	0.41
	(0.43)	(0.46)	(0.49)
	[499]	[869]	[532]
control variables			
age	43.09	44.94	49.10
	(10.74)	(14.72)	(14.85)
	[498]	[870]	[533]
age_square	1,971.92	2,236.17	2,630.77
	(994.55)	(1,418.62)	(1,495.30)
	[498]	[870]	[533]
expenditures_pcap	97.91	17.82	18.30
	(90.36)	(19.32)	(20.09)
	[499]	[870]	[532]
family_size	4.89	4.27	4.01
	(2.00)	(2.29)	(2.33)
	[499]	[870]	[533]
literate	0.53	0.28	0.33
	(0.50)	(0.45)	(0.47)
	[497]	[870]	[533]
married	0.87	-	-
	(0.33)	-	-
	[498]	-	-
member_socio	0.73	-	-
	(0.45)	-	-
	[498]	-	-

Notes: We report means, standard deviations (in parentheses), and sample sizes [in brackets]. The samples used to compute descriptive statistics for the ERHS survey data are heads aged 20 to 80 from the regions of Amhara and Oromia or heads and spouses aged 20 to 80 from the regions of Amhara and Oromia.

Table 2: **Observable characteristics (experimental and census data)**

	Experimental Data	Census Data	
		heads only	heads and spouses
age	43.09	41.94	38.84
	(10.74)	(15.08)	(14.45)
	[498]	[171534]	[283165]
age_square	1,971.92	1,986.44	1,717.84
	(994.55)	(1,426.06)	(1,313.05)
	[498]	[171,534]	[283,165]
family_size	4.89	4.66	4.97
	(2.00)	(2.25)	(2.19)
	[499]	[171,534]	[283,165]
female	0.10	0.23	0.53
	(0.30)	(0.42)	(0.50)
	[498]	[171,534]	[283,165]
literate	0.56	0.34	0.27
	(0.50)	(0.47)	(0.44)
	[497]	[171,534]	[283,165]
married	0.87	0.77	0.86
	(0.33)	(0.42)	(0.35)
	[498]	[171,534]	[283,165]

Notes: We report means, standard deviations (in parentheses), and sample sizes [in brackets]. The samples used to compute descriptive statistics for the census data includes heads aged 20 to 80 from the regions of Amhara and Oromia or heads and spouses aged 20 to 80 from the regions of Amhara and Oromia.

Table 3: **Housework, gender, and culture (experimental data)**

	(1) *house*	(2) *house*	(3) *house*	(4) *house*
female	0.4114***	0.3537**	0.3772**	0.3257**
	(0.1218)	(0.1375)	(0.1355)	(0.1511)
plow_pos	-0.0136	-0.0065	-0.0091	-0.0055
	(0.0183)	(0.0174)	(0.0149)	(0.0149)
female×plow_pos	0.1917	0.2116	0.2603	0.2770*
	(0.1592)	(0.1522)	(0.1633)	(0.1592)
c1	0.0150	0.0112	0.0056	0.0012
	(0.0123)	(0.0141)	(0.0107)	(0.0129)
c3	0.0139	0.0095	0.0081	0.0038
	(0.0159)	(0.0166)	(0.0158)	(0.0163)
c4	-0.0038	-0.0046	-0.0006	-0.0022
	(0.0194)	(0.0200)	(0.0152)	(0.0168)
age		0.0091*		0.0083*
		(0.0052)		(0.0048)
age_square		-0.0001		-0.0001
		(0.0001)		(0.0000)
married		-0.0305		-0.0246
		(0.0786)		(0.0667)
family_size		-0.0007		-0.0013
		(0.0040)		(0.0036)
literate		-0.0286*		-0.0209*
		(0.0147)		(0.0119)
member_socio		-0.0625**		-0.0538**
		(0.0225)		(0.0221)
ln_expenditures_pcap		-0.0014		-0.0011
		(0.0147)		(0.0154)
N	498	497	498	497
R^2	0.479	0.503	0.507	0.525
adj. R^2	0.473	0.490	0.501	0.513

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the session level). Regressions in specifications (3) and (4) include weights to account for unequal selection probabilities of households.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: **Housework, gender, and culture (ERHS)**

	(1) *house*	(2) *house*	(3) *house*	(4) *house*
	heads and spouses		heads only	
female	0.8068***	0.7791***	0.7273***	0.7316***
	(0.0481)	(0.0504)	(0.0532)	(0.0459)
plow_pos	0.0029	0.0025	0.0029	0.0025
	(0.0029)	(0.0075)	(0.0030)	(0.0075)
female×plow_pos	0.1670***	0.1264**	0.1929**	0.1908**
	(0.0490)	(0.0419)	(0.0612)	(0.0625)
head		-0.0753*		
		(0.0392)		
family_size		0.0053		0.0047
		(0.0035)		(0.0050)
age		-0.0001		-0.0020
		(0.0044)		(0.0047)
age_square		-0.0000		0.0000
		(0.0000)		(0.0001)
literate		-0.0190		-0.0236**
		(0.0129)		(0.0085)
ln_expenditures_pcap		-0.0025		-0.0066
		(0.0044)		(0.0081)
N	869	865	532	530
R^2	0.892	0.896	0.808	0.811
adj. R^2	0.891	0.895	0.807	0.808

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the *kebele* level). The samples include heads aged 20 to 80 from the regions of Amhara and Oromia or heads and spouses aged 20 to 80 from the regions of Amhara and Oromia.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: **Time-use, gender, and culture (ERHS, heads and spouses)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	housework		*childcare*		*livestock*		*farming*		*marketing*		*owntime*	
female	4.0363***	3.8039***	0.1313	0.1328	-0.0313	0.1746	-2.0131***	-1.9380***	0.0294	0.0731	-1.0892**	-1.1420***
	(0.5313)	(0.4387)	(0.1111)	(0.1346)	(0.1806)	(0.1849)	(0.4335)	(0.3905)	(0.1107)	(0.1077)	(0.3856)	(0.3392)
plow_pos	-1.0690**	-1.0761**	-0.1207	-0.1357	1.2342***	1.1455***	1.4033**	1.3473*	0.2360	0.2111	-0.3871*	-0.3561*
	(0.3341)	(0.3600)	(0.1324)	(0.1523)	(0.2811)	(0.2655)	(0.5942)	(0.6112)	(0.2045)	(0.2054)	(0.1972)	(0.1721)
female × plow_pos	1.2075**	0.8877*	0.2372***	0.1830**	-0.8930**	-0.7178*	-1.0378*	-1.0178*	-0.1792	-0.1698	0.2925	0.3571
	(0.3924)	(0.4636)	(0.0670)	(0.0682)	(0.3645)	(0.3478)	(0.4820)	(0.5251)	(0.2043)	(0.1950)	(0.2876)	(0.3389)
head		-0.6160		-0.0104		0.4162*		0.3502		0.0421		-0.3139
		(0.4905)		(0.1112)		(0.1946)		(0.2201)		(0.1093)		(0.1833)
family_size		0.0526		0.0251		0.0600*		0.0338		0.0093		-0.0610*
		(0.0511)		(0.0307)		(0.0318)		(0.0600)		(0.0226)		(0.0293)
age		0.0303		-0.0214		-0.0204		0.0320		-0.0114		-0.0073
		(0.0377)		(0.0143)		(0.0372)		(0.0277)		(0.0260)		(0.0371)
age_square		-0.0004		0.0002		0.0003		-0.0006*		0.0001		0.0004
		(0.0003)		(0.0002)		(0.0004)		(0.0003)		(0.0003)		(0.0004)
literate		-0.2591		-0.0113		0.0937		-0.0166		0.1377		0.1562
		(0.2721)		(0.1118)		(0.2181)		(0.4203)		(0.1616)		(0.1725)
ln_expenditures_pcap		0.0390		0.0599		0.0812		0.0271		-0.0248		0.0175
		(0.0713)		(0.0393)		(0.1011)		(0.1709)		(0.0418)		(0.0721)
N	869	865	869	865	869	865	869	865	869	865	869	865
R^2	0.467	0.472	0.023	0.030	0.069	0.080	0.187	0.197	0.004	0.007	0.034	0.058
adj. R^2	0.465	0.467	0.020	0.020	0.066	0.070	0.184	0.188	0.000	-0.004	0.030	0.048

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the *kebele* level). The sample includes heads and spouses aged 20 to 80 from the regions of Amhara and Oromia.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: **Time-use, gender, and culture (ERHS, heads only)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	housework		*childcare*		*livestock*		*farming*		*marketing*		*owntime*	
female	3.8322***	3.8451***	0.2606*	0.2755	0.0720	0.1728	-2.0061***	-2.0244***	0.1322	0.1903	-0.7465	-0.8832**
	(0.4933)	(0.5106)	(0.1275)	(0.1533)	(0.2099)	(0.2089)	(0.4417)	(0.4337)	(0.1167)	(0.1289)	(0.4094)	(0.3720)
plow_pos	-0.8888*	-0.9288*	-0.0279	-0.0381	1.2180***	1.1378***	1.3247*	1.2732*	0.2329	0.2067	-0.4247*	-0.3165
	(0.4127)	(0.4297)	(0.0799)	(0.1055)	(0.2884)	(0.2623)	(0.6121)	(0.6031)	(0.2100)	(0.2063)	(0.2033)	(0.1753)
female × plow_pos	0.9096*	0.9479*	0.0295	0.0151	-0.6618	-0.6858	-1.3142**	-1.1736*	-0.4042**	-0.3906*	-0.1685	-0.2674
	(0.4768)	(0.5017)	(0.1867)	(0.1856)	(0.3930)	(0.3708)	(0.5108)	(0.5516)	(0.1649)	(0.1722)	(0.3005)	(0.2839)
family_size		0.0318		0.0215		0.0771		0.0048		0.0142		-0.0952**
		(0.0715)		(0.0328)		(0.0467)		(0.0876)		(0.0359)		(0.0353)
age		0.0492		-0.0053		-0.0239		0.0930		-0.0222		-0.0466
		(0.0578)		(0.0129)		(0.0540)		(0.0606)		(0.0376)		(0.0379)
age_square		-0.0005		0.0000		0.0003		-0.0012*		0.0002		0.0007*
		(0.0005)		(0.0001)		(0.0006)		(0.0006)		(0.0004)		(0.0004)
literate		-0.0275		-0.0627		0.0186		-0.0096		0.1422		-0.0598
		(0.2980)		(0.0550)		(0.2533)		(0.5562)		(0.1841)		(0.1446)
ln_expenditures_pcap		0.0797		0.0515		0.1373		-0.0229		-0.0236		0.0049
		(0.0925)		(0.0561)		(0.1566)		(0.1991)		(0.0705)		(0.0527)
N	532	530	532	530	532	530	532	530	532	530	532	530
R^2	0.431	0.435	0.021	0.028	0.045	0.056	0.138	0.148	0.006	0.009	0.029	0.069
adj. R^2	0.428	0.426	0.015	0.013	0.039	0.041	0.133	0.135	-0.000	-0.006	0.023	0.054

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the *kebele* level). The sample includes heads aged 20 to 80 from the regions of Amhara and Oromia.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: **Experimental outcome, gender, and culture (experimental data)**

	(1) *log grams*	(2) *log grams*	(3) *log grams*	(4) *log grams*	(5) *log grams*	(6) *log grams*
female	0.1613***	0.0801*	0.0513	0.1631***	0.0645	0.0352
	(0.0211)	(0.0435)	(0.0682)	(0.0257)	(0.0492)	(0.0761)
c1	0.0244	0.0229	0.0205	0.0211	0.0191	0.0171
	(0.0555)	(0.0551)	(0.0553)	(0.0608)	(0.0599)	(0.0609)
c3	0.1104*	0.1124**	0.1243**	0.1415**	0.1411**	0.1505***
	(0.0548)	(0.0537)	(0.0545)	(0.0508)	(0.0515)	(0.0522)
c4	0.0920*	0.0927*	0.0925*	0.0669	0.0658	0.0677
	(0.0474)	(0.0511)	(0.0479)	(0.0467)	(0.0481)	(0.0469)
plow_pos		-0.0124	0.0051		-0.0275	-0.0096
		(0.0532)	(0.0520)		(0.0494)	(0.0503)
female×plow_pos		0.1352**	0.1376**		0.1463**	0.1519*
		(0.0541)	(0.0633)		(0.0642)	(0.0750)
age			0.0088			0.0085
			(0.0060)			(0.0057)
age_square			-0.0002**			-0.0002**
			(0.0001)			(0.0001)
married			-0.0225			-0.0159
			(0.0370)			(0.0394)
family_size			0.0028			0.0015
			(0.0044)			(0.0041)
literate			-0.0019			-0.0092
			(0.0190)			(0.0179)
member_socio			-0.0347			-0.0270
			(0.0207)			(0.0235)
ln_expenditures_pcap			-0.0048			-0.0110
			(0.0136)			(0.0148)
N	498	498	497	498	498	497
R^2	0.124	0.133	0.236	0.137	0.146	0.242
adj. R^2	0.117	0.123	0.215	0.130	0.135	0.222

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the session level). Regressions in specifications (4), (5) and (6) include weights to account for unequal selection probabilities of households.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: **Housework, gender, and transmission of culture (experimental data)**

	(1) *house*	(2) *house*	(3) *house*	(4) *house*
plow_pos_keb	-0.0222	0.0002	-0.0196	-0.0034
	(0.0220)	(0.0199)	(0.0202)	(0.0185)
plow_pos	-0.0161	-0.0129	-0.0149	-0.0119
	(0.0221)	(0.0232)	(0.0190)	(0.0197)
plow_pos×plow_pos_keb	-0.0171	-0.0052	-0.0142	-0.0064
	(0.0211)	(0.0196)	(0.0191)	(0.0184)
female	0.5152***	0.4660***	0.5691***	0.5266***
	(0.1506)	(0.1569)	(0.1660)	(0.1682)
female×plow_pos_keb	0.1410	0.0978	0.0963	0.0579
	(0.1640)	(0.1893)	(0.1243)	(0.1487)
female×plow_pos	0.4125**	0.3980*	0.4488**	0.4378**
	(0.1913)	(0.2077)	(0.1742)	(0.1867)
female×plow_pos×plow_pos_keb	0.6438***	0.6186***	0.6617***	0.6374***
	(0.0897)	(0.1140)	(0.0890)	(0.1098)
c1	0.0157	0.0136	0.0072	0.0041
	(0.0130)	(0.0138)	(0.0111)	(0.0126)
c3	0.0141	0.0127	0.0080	0.0056
	(0.0175)	(0.0184)	(0.0157)	(0.0167)
c4	-0.0061	-0.0046	-0.0064	-0.0072
	(0.0182)	(0.0196)	(0.0142)	(0.0163)
age		0.0084*		0.0077*
		(0.0046)		(0.0042)
age_square		-0.0001*		-0.0001*
		(0.0000)		(0.0000)
married		-0.0215		-0.0152
		(0.0737)		(0.0647)
family_size		-0.0023		-0.0031
		(0.0039)		(0.0034)
literate		-0.0259*		-0.0185
		(0.0143)		(0.0115)
member_socio		-0.0630***		-0.0519**
		(0.0214)		(0.0197)
ln_expenditures_pcap		0.0029		0.0040
		(0.0135)		(0.0139)
N	498	497	498	497
R^2	0.512	0.535	0.551	0.567
adj. R^2	0.502	0.518	0.542	0.552

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the session level). Regressions in specifications (3) and (4) include weights to account for unequal selection probabilities of households.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: **Housework, gender, and transmission of culture (ERHS)**

	(1) *house* heads and spouses	(2) *house* heads and spouses	(3) *house* heads only	(4) *house* heads only
plow_pos_keb	-0.0000	0.0143	0.0000	0.0147**
	(0.0000)	(0.0125)	(.)	(0.0048)
plow_pos	-0.0000	0.0005	0.0000	0.0006
	(0.0000)	(0.0069)	(.)	(0.0061)
plow_pos×plow_pos_keb	0.0083	0.0068	0.0110	0.0114
	(0.0070)	(0.0119)	(0.0060)	(0.0113)
female	0.8000***	0.7716***	0.6818***	0.6896***
	(0.0491)	(0.0511)	(0.0526)	(0.0479)
female×plow_pos_keb	1.0000***	0.9739***	0.9091***	0.8958***
	(0.0000)	(0.0258)	(0.0216)	(0.0215)
female×plow_pos	0.9762***	0.9047***	0.8571***	0.8595***
	(0.0110)	(0.0427)	(0.0686)	(0.0723)
female×plow_pos×plow_pos_keb	0.9776***	0.9643***	0.9670***	0.9501***
	(0.0194)	(0.0400)	(0.0369)	(0.0334)
head		-0.0758*		
		(0.0397)		
family_size		0.0048		0.0032
		(0.0035)		(0.0048)
age		-0.0000		-0.0006
		(0.0045)		(0.0045)
age_square		-0.0000		-0.0000
		(0.0000)		(0.0000)
literate		-0.0205		-0.0221**
		(0.0135)		(0.0080)
ln_expenditures_pcap		-0.0026		-0.0055
		(0.0036)		(0.0074)
N	869	865	532	530
R^2	0.892	0.897	0.815	0.817
adj. R^2	0.891	0.895	0.813	0.813

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the *kebele* level). The samples include heads aged 20 to 80 from the regions of Amhara and Oromia or heads and spouses aged 20 to 80 from the regions of Amhara and Oromia.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: **Experimental outcome, gender, and transmission of culture (experimental data)**

	(1) *log grams*	(2) *log grams*	(3) *log grams*	(4) *log grams*
plow_pos_keb	-0.0722	-0.0931	-0.0499	-0.0572
	(0.1104)	(0.1391)	(0.1085)	(0.1338)
plow_pos	0.0836	0.0750	0.0897*	0.0817
	(0.0511)	(0.0505)	(0.0517)	(0.0514)
plow_pos×plow_pos_keb	-0.0475	-0.0293	-0.0548	-0.0364
	(0.0655)	(0.0631)	(0.0611)	(0.0596)
female	0.0988***	0.0796	0.0940***	0.0720
	(0.0257)	(0.0526)	(0.0192)	(0.0502)
female×plow_pos_keb	-0.0048	-0.0328	-0.0057	-0.0311
	(0.1014)	(0.1130)	(0.0916)	(0.1047)
female×plow_pos	0.2961***	0.2836***	0.2957***	0.2888***
	(0.0673)	(0.0724)	(0.0651)	(0.0713)
female×plow_pos×plow_pos_keb	0.1478**	0.1467*	0.1377*	0.1376*
	(0.0680)	(0.0723)	(0.0677)	(0.0694)
c1	0.0106	0.0121	0.0118	0.0118
	(0.0535)	(0.0544)	(0.0598)	(0.0609)
c3	0.0922	0.1062*	0.1259**	0.1367**
	(0.0540)	(0.0547)	(0.0525)	(0.0528)
c4	0.0486	0.0553	0.0358	0.0415
	(0.0429)	(0.0421)	(0.0395)	(0.0397)
age		0.0080		0.0079
		(0.0060)		(0.0057)
age_square		-0.0001**		-0.0001**
		(0.0001)		(0.0001)
married		-0.0164		-0.0105
		(0.0404)		(0.0422)
family_size		0.0023		0.0010
		(0.0044)		(0.0041)
literate		0.0049		-0.0050
		(0.0180)		(0.0180)
member_socio		-0.0315		-0.0250
		(0.0203)		(0.0232)
ln_expenditures_pcap		-0.0046		-0.0119
		(0.0137)		(0.0145)
N	498	497	498	497
R^2	0.173	0.264	0.180	0.267
adj. R^2	0.156	0.238	0.163	0.241

Notes: All regressions include a constant. Standard errors are reported in parentheses (clustered at the session level). Regressions in specifications (4), (5) and (6) include weights to account for unequal selection probabilities of households.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix - For Online Publication

A Definition of Variables

A.1 Experimental Data

Variable	Description
age	age of the participant, in years
age_square	square of *age*
c1	binary variable equal to one if the session was assigned to the wage treatment where $m = 20$ Birr, zero otherwise
c2	binary variable equal to one if the session was assigned to the sharecropping treatment where $m = grams \times 0.2$ Birr, zero otherwise
c3	binary variable equal to one if the session was assigned to the fixed-rent treatment where $m = grams \times 0.4$ Birr -20 Birr, zero otherwise
c4	binary variable equal to one if the session was assigned to the ownership treatment where $m = grams \times 0.4$ Birr, zero otherwise
expenditures_pcap	average weekly per-capita food and non-food expenditures (including expenditure for energy and transportation, education, health, social events, livestock, housing, and other non-food expenditures), in Birr, computed following Deaton (1997): Children between 0 and 5 years of age receive a weight of 0.4 adults, children between 6 and 11 receive a weight of 0.5 adults, household members of age 12 and older are considered adults and receive a weight of 1
family_size	number of persons living in the participant's family (excluding the participant)
female	binary variable equal to one if the participant is female, zero if he is male
grams	output achieved by the participant in the experimental task, i.e., the weight (measured in grams) of the beans sorted from the blue into the red bucket within 15 minutes
house	binary variable equal to one if the participant's main occupation in the last 12 months was housework, zero otherwise
literate	binary variable equal to one if the participant ever received any schooling, zero otherwise
married	binary variable equal to one if the participant is married, zero otherwise
member_socio	binary variable equal to one if the participant is a member of a socio-political organization such as a cooperative, zero otherwise
plow_pos	binary variable equal to one if the participant's household cultivates at least one plow-positive crop (white teff, other teff, barley, or wheat), zero otherwise
plow_pos_keb	binary variable equal to one if at least 84 percent of the participants in a session cultivate at least one plow-positive crop, i.e., if the share of households that cultivate plow-positive crops is greater than or equal to the average share of households in the sample that cultivate plow-positive crops, zero otherwise

Notes: Data collection took place in April and May, 2011.

A.2 Population and Housing Census

Variable	Description
age	age of the respondent, in years
age_square	square of *age*
family_size	number of persons living in the respondent's home in 2007 (excluding the respondent)
female	binary variable equal to one if the respondent is female, zero if he is male
literate	binary variable equal to one if the respondent is literate, zero otherwise
married	binary variable equal to one if the respondent is married, zero otherwise

Notes: We use data from the Population and Housing Census 2007.

A.3 Ethiopian Rural Household Survey

Variable	Description
age	age of the respondent in 1997, in years [1994]
age_square	square of *age* [1994]
childcare	time the respondent allocated to childcare (including breastfeeding, other exclusive feeding, bathing/dressing, teaching, playing/comforting/watching children, and caring for sick children) during the previous day, in hours [1997]
expenditures_pcap	weekly per-capita food and non-food expenditures (including expenditures for transportation, social events, taxes, education, and health), in Birr, computed following Deaton (1997): Children between 0 and 5 years of age receive a weight of 0.4 adults, children between 6 and 11 receive a weight of 0.5 adults, household members of age 12 and older are considered adults and receive a weight of 1 [1997]
farming	time allocated to working in farm land preparation/planting/weeding, harvesting, etc., and fencing during the previous day, in hours [1997]
female	binary variable equal to one if the respondent is female, zero if he is male [1994]
head	binary variable equal to one if the respondent is the household head, i.e., the member of the household who is recognized as the head by the other household members, zero otherwise [1994]
family_size	number of persons living in the respondent's household in 1994 (excluding the respondent) [1994]
house	binary variable equal to one if the respondent's main activity is housework, zero otherwise [1997]
housework	time the respondent allocated to domestic work (including cooking/meal preparation/washing dishes, coffee preparation, cleaning house/paddock, fetching water, gathering firewood, and dungcake preparation) during the previous day, in hours [1997]
literate	binary variable equal to one if the respondent is literate, i.e., knows how to read and write, zero otherwise [1994]
livestock	time the respondent allocated to tending livestock, milking cows, and herding during the previous day, in hours [1997]
marketing	time the respondent allocated to buying/selling of crops, and other marketing during the previous day, in hours [1997]
owntime	time the respondent allocated to eating, resting/chatting/relaxing, praying, and bathing/personal care during the previous day, in hours [1997]
plow_pos	binary variable equal to one if the respondent's household cultivates at least one plow-positive crop (white teff, other teff, barley, or wheat), zero otherwise [1997]
plow_pos_keb	binary variable equal to one if at least 85 (81 in the 'heads only' sample) percent of the respondents in a *kebele* cultivate at least one plow-positive crop, i.e., if the share of households that cultivate plow-positive crops is greater than or equal to the average share of households in the sample that cultivate plow-positive crops, zero otherwise [1997]

Notes: We used data from the Ethiopian Rural Household Survey, First Round (1994) and Fourth Round (1997). The first round was used to identify the households and their members. Socio-demographic information was also taken from the first round. The information on time-use, cultivated crops, and expenditures was taken from the fourth round. The years in brackets indicate which survey round was used to generate the variables.

B Sampling Procedure

The selection of 20 participating households in each of the 25 *kebeles* occurred in four stages:

I We defined the sample universe to consist of households in the Amhara and Oromia regions.

II We sampled two districts (*woredas*) in each region: Gozamen and Bahir Dar Zuria in Amhara, and Adaa and Girar Jarso in Oromia.

III Within these *woredas*, we randomly selected 25 *kebeles*: 8 in Gozamen, 7 in Bahir Dar Zuria, 5 in Adaa, and 5 in Girar Jarso.

IV Using registries kept by the *kebele* chiefs, we invited every $N/22$ household heads from the registry (20 participants and two replacements in case the sampled participants did not show up for the experiment). This was done by an enumerator who traveled to the study sites one day before the data collection.

We define the household head to be the person with the best knowledge of the household's decision making and the household's resources. We decided to invite only household heads rather than randomly selected household members to the experiment because it was accompanied by a detailed quantitative and qualitative household survey. Household heads were considered the most appropriate household members to answer the survey items.

This procedure generates a sample in which the households are not randomly distributed across the two regions, but geographically clustered in *kebeles*. Given the varying population sizes in the clusters, the households were sampled with unequal selection probabilities. To take this into account, we also present results from population-weighted regressions. This ensures that the estimated parameters are representative for the underlying population. The population weights were calculated as the inverse of the sum of the sampled fraction in each region, i.e., the population in sampled districts to the total population in all districts in the two regions, and the population in the sampled *kebeles* to the total population in all *kebeles* in the sampled districts (Deaton, 1997).

C Independence

As participants might have increased their effort if they had been watching the others exercise a high level of effort (Falk and Ichino, 2006; Mas and Moretti, 2009), our biggest concern was to establish the independence of the effort exercised by each participant, so as to ensure our measuring the pure gender difference, rather than peer effects. We designed the experimental setup so that it prevented participants from influencing each other.

The enumerators arranged that the participants sit in a circle with an adequate distance between each other. The experiment was conducted outside, on meadows, to allow enough space. To prevent the participants from watching each other while performing the task, the enumerators asked them to turn around so that their backs were facing the inside of the circle. The enumerators stood inside the circle to ensure that the participants did not turn their heads to watch each other (see Figure 4 in Appendix E for a picture of the seating arrangements during the experiment). These measures were taken to avoid participants' adjusting their effort in reaction to watching how well others were performing the task. To also ensure that participants could not listen to each other performing the task, the bottoms of the red buckets were covered with a paper towel so that the beans did not make a sound when they were dropped into the buckets. Moreover, the enumerators played music on portable speakers while the participants were conducting the experimental task. The same song was played twice during each session. This was done to further avoid participants' adjusting their effort in reaction to hearing how well others were performing in the task. However, since the paper towel worked well, the music entertained participants more than it prevented them from hearing each other. After the seating was in order, the enumerators arranged the buckets. They asked each participant which hand(s) they used to pick up things other than food. For right handers, the empty red bucket was put to their left, and for left handers, it was put to their right. The flat blue bucket filled with beans was put in front of their strong hand. The participants were then allowed to move the buckets to suit their own convenience, as long as they did not violate the basic setup. Time was called with a stopwatch, which was the same for all sessions. It was checked that all participants started and finished at the same time.

D Compensation Schemes

To generate the experimental output, we used the contractual relations commonly found in the land and labor markets in rural Ethiopia. There are four cases in the decision on how to contract for labor, land, and the claim to yields from land: wage labor, sharecropping, fixed-rent land leases, and ownership cultivation. These contracts were randomly assigned to sessions in the sampled *kebeles*.

In the wage treatment, the participants were compensated with the fixed amount of 20 Birr for performing the experimental task:

$$m = 20\ Birr \tag{5}$$

In the sharecropping treatment, the compensation of participants was a fraction of the experimental output. As an equal split of the output between a landlord and a tenant is commonly observed in real sharecropping arrangements, the participants received one-half of 4 Birr for each additional 10 grams of experimental output:

$$m = grams \times 0.2\ Birr \tag{6}$$

where $grams$ is the output achieved in the experimental task.
In the fixed-rent treatment, the compensation of participants according to their experimental output was reduced by the fixed amount of 20 Birr:

$$m = grams \times 0.4\ Birr - 20\ Birr \tag{7}$$

In the ownership treatment, the participants were compensated according to their experimental output without any deduction:

$$m = grams \times 0.4\ Birr \tag{8}$$

On average, participants earned 20 Birr under the wage treatment, 26.43 Birr under the sharecropping treatment, 39.78 Birr under the fixed-rent treatment, and 58.43 Birr under the ownership treatment. These are sizable amounts of money, given that the average monthly per-capita expenditure of our subjects is

727 Birr.[31] For more details on the use of these contracts in the experiment, see Kemper and Schumacher (2014).

[31] These compensation schemes were not chosen arbitrarily. During preparatory sessions with the enumerators and students, the wage treatment yielded an average output of between 110 and 115 g. Before the start of the actual field work, we normalized the pay-off of the wage contract, sharecropping, and the fixed-rent contract, to be clearly below that value to avoid inappropriate scaling of the wage contract to influence the experimental output. That is, we normalized the pay-off of the wage contract, sharecropping, and the fixed-rent contract at a low level of experimental output: For $grams$ = 100 g, the pay-off was 20 Birr for each of the three treatments. The pay-off under the owner treatment was 40 Birr.

E Figures

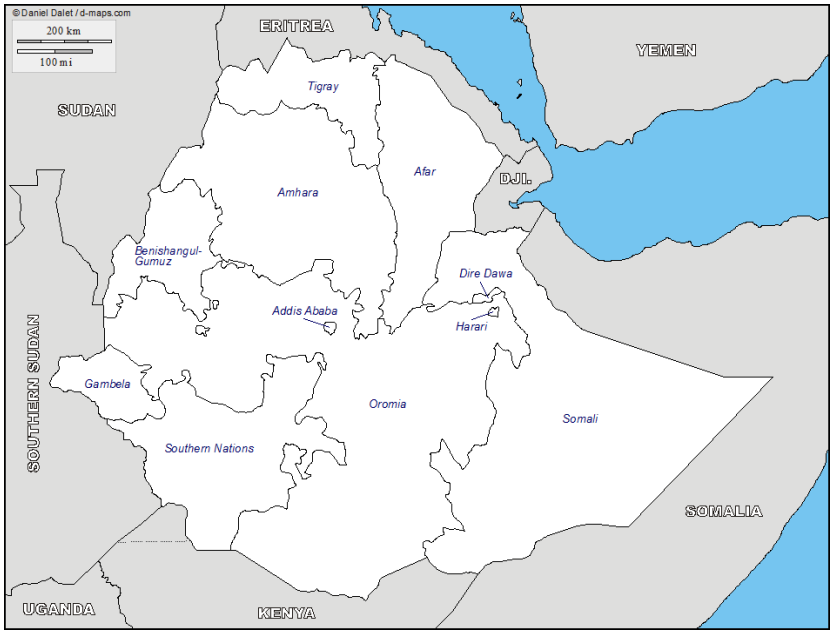

Figure 1: **Ethiopia.**
Source: `http://d-maps.com/carte.php?num_car=4258&lang=en`, retrieved on 15 January 2014.

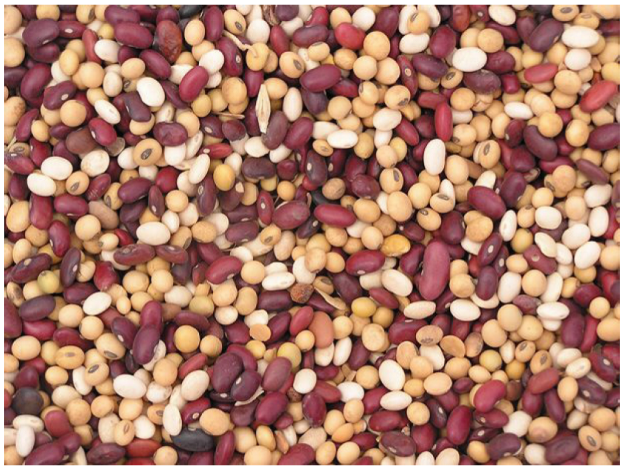

Figure 2: **Red, yellow and white beans.** For the experimental task, participants were asked to pick as many beans of a particular color as possible within a given period of time.

Figure 3: **Experimental equipment.** Each participant received a flat, blue bucket filled with 2.1 kg of beans in three different colors (700 g of each color), and an empty red bucket (without cover—unlike in the picture). Participants picked beans of one particular color from the blue bucket and dropped them into the red bucket. After time was up, the beans from the red bucket were weighed using a transparent bowl and a scale.

Figure 4: **Seating arrangements during the experiment.** To allow enough space, the experiment took place outside. Participants sat on the ground in a circle with their backs facing the inside of the circle to ensure that they could not see each other performing the task. Enumerators stood inside the circle.